科学版研究生教学丛书

应用泛函分析

胡适耕　编著

科学出版社
北京

内 容 简 介

本书介绍抽象空间、线性算子与线性泛函、谱论、非线性算子与非线性泛函的基本理论和基本方法，通过典型例题说明泛函分析方法在方程问题、逼近与计算理论、随机现象研究、工程与经济模型等诸多领域的应用.

本书可作为理工科诸专业高年级本科生、研究生的教学用书，也适合科技工作者和教师参考使用.

图书在版编目(CIP)数据

应用泛函分析/胡适耕编著.—北京:科学出版社,2003.8
(科学版研究生教学丛书)
ISBN 7-03-011664-X

Ⅰ.应… Ⅱ.胡… Ⅲ.泛函分析-研究生-教材 Ⅳ.O177

中国版本图书馆 CIP 数据核字(2003)第 058510 号

责任编辑:吕 虹 李鹏奇/责任校对:陈丽珠
责任印制:安春生/封面设计:黄华斌

科学出版社 出版
北京东黄城根北街16号
邮政编码:100717
http://www.sciencep.com

双青印刷厂 印刷

科学出版社发行 各地新华书店经销

*

2003年8月第 一 版 开本:B5(720×1000)
2003年8月第一次印刷 印张:16
印数:1—3 000 字数:298 000

定价: 26.00 元

(如有印装质量问题,我社负责调换〈环伟〉)

前　　言

如果你是"应用数学"、"信息与计算科学"专业的大学生，或是要求有较好数理基础的理工科专业硕士生，你多半需要学习"应用泛函分析"课程，这就可能有机会读到本书了．按照传统习惯，在开卷之初，你似乎总会看到一番宣示：泛函分析之作用如何重大，其历史又如何辉煌．这些无疑都是事实．可惜，一个从未接触过泛函分析的读者，未必能被这番话语打动．据笔者的经验，很少有学生去细读那些洋洋洒洒的前言．读者更感兴趣的多半是：即将开始学习的这门课程好对付吗？它与已经学过的数学课程，如微积分学、线性代数、微分方程等，有何异同？当然，关于这类问题，也未必能从简短的前言中找到完全的答案．不过，确有一些话，应该有言在先．

首先，不妨如实告诉你：与你已学过的数学课程比较，你很可能会感到泛函分析要难些．其缘由主要是，泛函分析具有更大的综合性；而与高度的综合性相适应的，自然是一定的抽象性，但很少有人会本能地喜欢抽象．例如，你在线性代数中已学过线性方程组，在微积分学中学过隐函数方程，在微分方程论中学过常微分方程与偏微分方程，等等．这些方程纵然都不容易研究，但似乎还可捉摸．在泛函分析中，所有这些方程被熔于一炉了，抽象成了所谓算子方程．再如，你在微积分学、最优化及变分学中，学习到在各种条件下求极值．这些问题背景清楚，实际意义明确．而到了泛函分析中，它们全被概括到"泛函极值"这一抽象概念中了．这就失去了问题的原型性与直观性，而这正是初学者难以适应的．然而，需要提醒你，这是要点，付出这一代价非常值得！惟有如此，才能实现问题的彻底简化，从而为便捷地求解开辟道路．例如，统一多种方程的算子方程 $x = Fx$ 形式简单至极，在抽象空间的框架下，更便于对它作出深刻的分析．而且，通过类比与联想，在抽象空间中表述的问题亦呈现出某种直观的面貌，这就使在抽象化过程中失去直观性的问题，在更高层次上重新获得了直观，而且往往是更易于洞察与把握的直观．初看起来，真有些不可思议：经过抽象处理的问题往往更直观！你从泛函分析的学习中，就将体会到这种奥妙．

同样重要的问题是，你关心泛函分析的应用．很多泛函分析书籍都郑重宣告：泛函分析将应用到物理学、信息科学、工程技术乃至经济学．这些全是事实．然而，如果你指望从一本泛函分析书籍中看到所有这些应用，那么你多半会感到失望．实际上，这出于对一个数学学科的误解，我们将在此处尽量澄清这种误解．当将数学应用于某一实际问题时，首要的事情是建立适当的数学模型；而用于表述模型的数

学工具可能是微分方程、最优化或者随机过程等，很少直接运用更高层次的泛函分析概念来建模. 在更多的情况下，是将泛函分析理论用于模型分析. 当然，这最终为解决实际问题提供了强有力的方法. 但很明显，泛函分析的应用既离不开问题所涉及的专业知识，也离不开用于建模的那些数学知识(如微分方程、最优化等). 这就决定了泛函分析的应用很不简单直接，真能用作说明的例子可没有办法随手拈来. 且不说作者自身不可避免地受到其专业知识的局限，即使真正联系于某一专门领域(例如经济学)的应用，亦未必能使其他专业的读者感兴趣. 不妨强调一下，本书不是"泛函分析的应用"，而是"应用泛函分析"；对于后者，可解读为"供应用的泛函分析"，它提供泛函分析的基本概念、结论与方法，并为一些潜在的应用尽可能地指明方向. 至于关于其应用的具体与深入的讨论，则是各个专门学科的任务，绝非泛函分析一家所能独揽的. 如果说，"应用泛函分析"毕竟有别于"泛函分析"的话，那只是前者更强调泛函分析的应用背景，更注重泛函分析与邻近学科的联系，更重视其主要方法的实际可操作性. 在这些方面，本书力求做得好一些，其效果如何，只能由读者来评判.

作者希望本书能为不同类型的学校与不同需要的读者所用——从仅需最基本的材料到需要稍深入的知识之间的几个层次. 这一意图决定了本书的选材与结构. 使用本书作为教材时，可依几种方式组合书中的材料. 若选用§1.1～§1.5, §2.1～§2.5, §3.1～§3.4,则需 48 学时左右；若加上§3.5, §3.1～§3.3,则需 60 学时左右；讲完全部内容，约要 70 学时. 关于内容的取舍，使用本书的教师是最权威的. 他们富有创意的运用，将是对作者个人经验极有益的补充与修正，这正是作者所期待的.

胡适耕

记号与约定

A^c：集 A 的补.

A°：集 A 的内部.

$\overline{A}$：集 A 的闭包.

$A^\perp$：集 A 的正交补或零化子.

$B_r(a) = B(a,r)$：以 a 为心以 r 为半径的开球.

$\overline{B}_r(a) = \overline{B}(a,r)$：以 a 为心以 r 为半径的闭球.

$B(\Omega)$：Ω 上的有界函数之集.

$\mathbf{C}$：复数域.

$C(A,B)$：从 A 到 B 的连续映射之全体；$C(A) = C(A,\mathbf{K})$.

$C^m(\Omega)$：Ω 上的 m 次连续可微函数之全体；$C(\Omega) = C^0(\Omega)$.

$\mathrm{CL}(X,Y)$：从 X 到 Y 的紧线性算子之全体；$\mathrm{CL}(X) = \mathrm{CL}(X,X)$.

$D(T)$：算子 T 的定义域.

$d(A,B)$：集 A 与 B 之间的距离；$d(x,B) = d(\{x\},B)$.

$\mathrm{diam}A$：集 A 的直径.

$\mathrm{dim}X$：空间 X 的维数.

δ_{ij}：Kronecker 记号.

∂A：集 A 的边界.

$\{e_i\}$：未加说明时表 $\mathbf{R}^n$ 或 l^p 的标准基.

F：通常表示某个映射或算子；f：通常表示泛函.

$\mathrm{GL}(X)$：X 上拓扑自同构之全体.

H：通常记 Hilbert 空间.

I：单位算子.

J：常用来表示区间 $[a,b](a < b)$.

$\mathbf{K} = \mathbf{R}$ 或 $\mathbf{C}$；$\mathbf{K}^n$：n 维 Euclid 空间.

$\mathbf{K}^{m\times n}$：$\mathbf{K}$ 上 $m \times n$ 阶矩阵之全体；$A \in \mathbf{K}^{m\times n}$ 常写作 $[a_{ij}]$ 或 $[a_{ij}]_{m\times n}$.

L^p：p 次可积函数空间，$1 \leqslant p < \infty$.

L^∞：本性有界函数空间.

$L(X,Y)$：从 X 到 Y 的有界线性算子之全体；$L(X) = L(X,X)$.

m：Lebesgue 测度.

$\mathbf{N}$:自然数集.

$N(T)$: 算子 T 的零空间.

$\mathbf{Q}$:有理数集.

$\mathbf{R}$:实数集; $\mathbf{R}_+ = [0,\infty)$.

$R(F)$: 映射 F 的值域.

$R(\lambda,T)$ 或 R_λ: 算子 T 的预解式.

$r_\sigma(T)$: 算子 T 的谱半径.

$\rho(T)$: 算子 T 的正则值集.

S^1: 单位圆周.

$\mathrm{span}A$: 集 A 生成的向量空间.

$\mathrm{supp}f$: 函数 f 的支集.

$\sigma(T)$: 算子 T 的谱.

$\sigma_p(T)$: 算子 T 的点谱.

T: 通常表示线性算子; $T_\lambda = \lambda I - T$.

T^*: 算子 T 的对偶算子或相伴算子.

$\mathrm{V}_a^b(f)$: 函数 f 在 $[a,b]$ 上的全变差.

X,Y,Z: 通常记赋范空间.

X^*: 空间 X 的对偶空间.

$\mathbf{Z}$:整数集;$\mathbf{Z}_+$:非负整数集.

χ_A: 集 A 的特征函数.

$\triangleq$: 定义为.

$\equiv$:恒等于.

$\square$:定理或命题证完.

几点说明

1.指标用法 不致误解时,出现于符号 $\sum$, $\prod$, $\cup$, $\cap$ 下的指标予以省略.未加说明时,下标 n 遍取自然数.任给 $x \in \mathbf{R}^n$, $f: D \to \mathbf{R}^n$,自动认定 $x = (x_i) = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$, $f = (f_i) = (f_1, f_2, \cdots, f_n)^{\mathrm{T}}$.任给 $x \in l^p$,认定 $x = (x_i) = (x_1, x_2, \cdots, x_i, \cdots)$,涉及与无穷矩阵相乘时看做列向量.

2.极限符号 $\lim\limits_n x_n$ 与 $\lim\limits_{m,n} x_{mn}$ 分别表示 $\lim\limits_{n\to\infty} x_n$ 与 $\lim\limits_{m,n\to\infty} x_{mn}$; $\lim\limits_{t\downarrow 0}\varphi(t) = \lim\limits_{0<t\to 0}\varphi(t)$, $\rightrightarrows$ 表示一致收敛; $\xrightarrow{L^p}$ 表示 L^p 收敛; $\rightharpoonup$ 与 $\overset{*}{\rightharpoonup}$ 分别表示弱收敛与弱 * 收敛; $\overline{\lim\limits_n} x_n = \lim\limits_n \sup\limits_{k\geqslant n} x_k$, $\underline{\lim\limits_n} x_n = \lim\limits_n \inf\limits_{k\geqslant n} x_k$.

3.范数记号 不致混淆时,空间 X, Y, X^*, $L(X, Y)$ 等中的范数皆记作 $\|\cdot\|$; $\|\cdot\|_p$ 记 L^p 范数; $\|\cdot\|_0$ 记 sup 范数; $|\cdot|$ 记 $\mathbf{K}^n$ 中的 Euclid 范数.

4.零记号 数零、零向量、零算子与零泛函均记作 0.

5.集记号 集的标准表示是: $A = \{x : x$ 满足某条件$\}$,例如 $\{x : x \in \mathbf{R}$ 且 $f(x) > 0\}$,也写作 $\{x \in \mathbf{R} : f(x) > 0\}$ 或简写作 $\{f > 0\}$. $f(A) = \{f(a) : a \in A\}$; $x + A = \{x + a : a \in A\}$.

6.不等号的用法 对于 $A, B \subset \mathbf{R}$,约定 $A < B \Leftrightarrow \forall a \in A, \forall b \in B$,有 $a < b$; $A \leqslant B$, $a \leqslant B$ 等仿此.对 $f, g : D \to \mathbf{R}$,约定 $f < g \Leftrightarrow \forall x \in D$,有 $f(x) < g(x)$; $f \leqslant g$ 仿此; $f(A) < f(B) \Leftrightarrow \forall a \in A, \forall b \in B$,有 $f(a) < f(b)$; $f(A) \leqslant f(B)$ 仿此.

7.映射记号 $F : D \subset X \to Y$ 表示映射 $F : D \to Y$, D 是 X 的子集. $\varphi(x, y)$ 看作 x 的函数时记作 $\varphi(\cdot, y)$, $\cdot$ 表示自变量.若 $Fx \equiv x (\forall x \in X)$,则将 F 写作 I_X 或 I.

目　　录

第一章　抽 象 空 间

本书所讲述的泛函分析理论与方法，将在一种类似于平常空间但更一般的数学结构中展开. 这类结构以具有极限运算为其基本特征，通常统称为抽象空间. 我们将重点介绍最常用的 Banach 空间，同时也论及稍特殊的 Hilbert 空间与更一般的度量空间. 有关这些空间的基本知识是进一步学习泛函分析的基础.

§1.1　Banach 空间

§1.1.1　赋范空间

空间概念的推广，对于你可能早已不是什么新奇的事了. 实际上，当你在线性代数中学习向量空间时，已经实现了从直观的 3 维空间到抽象的 n 维空间的飞跃. 如你熟知，Euclid 空间 $\mathbf{R}^n$ 由所有形如 $x=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$（下面缩写作$(x_i)$）的 n 维向量组成，其中定义了线性运算：

$$\alpha(x_i)+\beta(y_i)=(\alpha x_i+\beta y_i),$$

α,β 是任意实数. 更重要的是，我们能将 3 维向量的长度公式自然地推广到 $\mathbf{R}^n$ 中：任给 $x=(x_i)\in\mathbf{R}^n$，称

$$|x|=\Big(\sum_i |x_i|^2\Big)^{1/2} \tag{1.1.1}$$

为向量 x 的**模长**. 从长度自然地引出极限概念：若 $\{x^{(k)}\}\subset\mathbf{R}^n$ 是一向量序列，$x\in\mathbf{R}^n$，则约定

$$x^{(k)}\to x \Leftrightarrow |x^{(k)}-x|\to 0 \quad (k\to\infty). \tag{1.1.2}$$

尽管 n 维 Euclid 空间已有一定的一般性，但还远不足以满足应用之需要. 现在面临一次更本质的跨越：从 n 维 Euclid 空间过渡到具有如下两方面结构的抽象空间（不妨记作 X）：

(i) 其中定义有线性运算；

(ii) 其中定义有“长度”，进而能由长度引出距离与极限.

为满足第一个要求，只需假定 X 是一般的向量空间①就够了. 至于“长度”，进一步

---

① 即线性空间，有关概念可参看任何一本线性代数教科书.

推广模长公式(1.1.1)似乎是一个自然的想法,但这至多适用于某些接近于 Euclid 空间的特殊系统,并不能达到一个一般的"长度"概念. 实际上,我们在利用模长公式(1.1.1)描述 $\mathbf{R}^n$ 中的极限运算时,很少需要直接运用表达式(1.1.1),而是利用模长的以下性质:

(i) $|\alpha x| = |\alpha||x|$;

(ii) $|x+y| \leqslant |x| + |y|$;

(iii) $|x| \geqslant 0$; $|x| = 0 \Leftrightarrow x = 0$($\mathbf{R}^n$ 的零元).

以上 $x, y \in \mathbf{R}^n$, $\alpha \in \mathbf{R}$. 这就启示出:若某个向量空间上定义了一种类似于模长的概念,它具有上述的性质(i)~(iii),则如同式(1.1.2)一样定义极限之后,就可将 Euclid 空间中那些仅依赖于性质(i)~(iii)的概念与结论推广于该空间. 以上想法引向赋范空间概念,它的严格定义在逻辑上是很简单的.

定义 1.1.1 设 X 是数域 $\mathbf{K}$(等于 $\mathbf{R}$ 或 $\mathbf{C}$,本书皆如此)上的向量空间. 若对每个 $x \in X$,指定了一个实数 $\|x\|$,称为 x 的**范数**[①],它满足以下**范数公理**:

(N_1)**齐次性**: $\|\alpha x\| = |\alpha| \|x\|$;

(N_2)**三角不等式**: $\|x+y\| \leqslant \|x\| + \|y\|$;

(N_3)**正定性**: $\|x\| \geqslant 0$; $\|x\| = 0 \Leftrightarrow x = 0$.

(以上 $x, y \in X$, $\alpha \in \mathbf{K}$),则称 X 为 $\mathbf{K}$ 上的赋范向量空间,简称为**赋范空间**;当必须明确指出范数时写作 $(X, \|\cdot\|)$. 若 $\mathbf{K} = \mathbf{R}$(或 $\mathbf{C}$),则称 $\mathbf{K}$ 上的赋范空间为实(或复)赋范空间.

依定义 1.1.1, $\mathbf{R}^n$ 就是一个实赋范空间,其中的范数就是依式(1.1.1)定义的模长,也称它为 **Euclid 范数**. 类似地,同一公式(1.1.1)定义的范数使 $\mathbf{C}^n$ 成为复赋范空间. 今后以 $\mathbf{K}^n$ 泛指 $\mathbf{R}^n$ 与 $\mathbf{C}^n$,它经常用作解释赋范空间概念的模型.

与特殊的 Euclid 范数 $|\cdot|$ 不同,定义 1.1.1 中的**一般的**范数 $\|\cdot\|$ 无具体计算公式. 对此,需要说明两点:首先,就理论分析而言,真正本质的东西恰恰是范数公理(N_1)~(N_3),而不是范数的具体计算公式;正因为舍弃了范数的特殊表达式,才得以建立具有高度一般性的赋范空间理论. 另一方面,一旦需要将赋范空间理论应用于某一特定空间,就必须选择适当的范数公式,并验证它确实满足范数公理(N_1)~(N_3). 现在就来看一个简单例子.

例 1.1.2(有界函数空间) 设 Ω 是任一非空集, $B(\Omega)$ 是定义于 Ω 上的有界实(或复)函数之全体,它显然是一个实(或复)向量空间. 任给 $u \in B(\Omega)$,令

$$\|u\|_0 = \sup_{x \in \Omega} |u(x)|, \tag{1.1.3}$$

① 范数一词,今后将兼指特定向量 x 的范数 $\|x\|$ 与函数 $x \to \|x\|$. 在后一种意义上使用时,通常写作 $\|\cdot\|$,其中·代表变量 x.

则可直接验证：

$$\begin{aligned}\|\alpha u\|_0 &= \sup_{x\in\Omega}|\alpha u(x)| \\ &= |\alpha|\sup_{x\in\Omega}|u(x)| = |\alpha|\,\|u\|_0,\\ \|u+v\|_0 &= \sup_{x\in\Omega}|u(x)+v(x)| \\ &\leqslant \sup_{x\in\Omega}|u(x)| + \sup_{x\in\Omega}|v(x)| \\ &= \|u\|_0 + \|v\|_0,\end{aligned}$$

$$\|u\|_0 \geqslant 0;\ \|u\|_0 = 0 \Leftrightarrow u(x)\equiv 0 \Leftrightarrow u = 0,$$

以上 $u,v\in B(\Omega)$，α 是常数．这就验证了范数公理$(N_1)\sim(N_3)$，因而 $\|\cdot\|_0$ 确是一范数，称为 **sup 范数或上确界范数**；$B(\Omega)$ 依 sup 范数是一赋范空间．今后使用空间 $B(\Omega)$ 时，总假定其中使用 sup 范数．

自然数集 $\mathbf{N}$ 上的有界函数就是有界数列．因此，有界数列空间 $B(\mathbf{N})$ 是一赋范空间，通常记作 l^∞ 或 m；对每个 $x=(x_i)\in l^\infty$，其 sup 范数为

$$\|x\|_0 = \sup_i |x_i|.$$

不过，在这一特殊情况下，通常记作 $\|x\|_\infty$．

在下节中，我们将系统地介绍多种具体的赋范空间．

以下设 X 是一个给定的($\mathbf{K}$ 上的)赋范空间．

为加强与平常 Euclid 空间的类比，我们将大量借用通常的几何术语，赋予抽象概念以某种直观形象．例如，X 中的元称为点或向量；向量 x 亦解释为从原点(即零元 0)到点 x 的有向线段(图 1-1)，而 $\|x\|$ 即其"长度"；当 $\|x\|=1$ 时称 x 为**单位向量**．经常用到的一个简单事实是：若 $0\neq x\in X$，则 $x/\|x\|$ 是单位向量．向量 y 也可表示为从点 x 到 $x+y$ 的有向线段．三角不等式无非是说：三角形一边之长不超过另两边长之和．如同在 $\mathbf{R}^3$ 中一样，由公理(N_2)易推出：

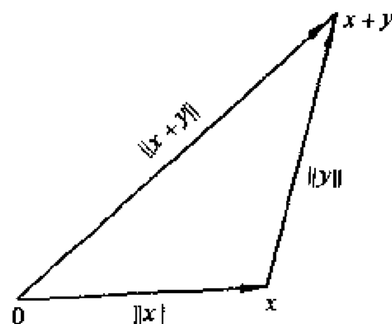


图 1-1

$$\left\|\sum_1^n x_i\right\| \leqslant \sum_1^n \|x_i\|, \tag{1.1.4}$$

$$\big|\|x\|-\|y\|\big| \leqslant \|x-y\|, \tag{1.1.5}$$

以上 $x_i,x,y\in X(1\leqslant i\leqslant n)$．你应当能说出这两个不等式的几何意义．对任给

$x,y\in X$，称 $\|x-y\|$ 为点 x 与 y 之间的距离，也记作 $d(x,y)$. 更一般地，对任给非空集 $A,B\subset X$，定义 A 与 B 之间的距离为

$$d(A,B)=\inf_{a\in A,b\in B}\|a-b\|. \tag{1.1.6}$$

约定 $d(x,B)=d(\{x\},B)$. 定义 A 的**直径**为

$$\operatorname{diam}A=\sup_{x,y\in A}\|x-y\|, \tag{1.1.7}$$

当 $\operatorname{diam}A<\infty$ 时称 A 为有界集. 显然，A 有界可等价地刻画为

$$\sup_{x\in A}\|x\|<\infty.$$

§1.1.2　极限与完备性

上面定义了赋范空间 X 中的距离，因而可描述空间中两点的接近程度，而这正是定义极限的基础. 形式上，为在 X 中定义极限，几乎只需套用 $\mathbf{R}^n$ 中的极限定义式(1.1.2).

定义 1.1.3　设 $\{x_n\}\subset X$ 是一序列，$x\in X$. 若

$$\|x_n-x\|\to 0\quad(n\to\infty),$$

则说序列 $\{x_n\}$ **范数收敛**(简称为收敛)于极限 x，记作

$$x_n\to x(n\to\infty)\quad 或\quad \lim_n x_n=x.$$

由定义直接看出，在 X 中 $x_n\to x\Leftrightarrow x_n-x\to 0(n\to\infty)$. 因此，要描述空间 X 中的范数收敛，只需解释在其中 $x_n\to 0$ 的意义就够了. 对于一个具体给定的赋范空间，确切地指明 $x_n\to 0$ 的具体含义，无疑是一件重要的事. 对于 Euclid 空间 $\mathbf{K}^n$ 中的序列

$$x^{(k)}=(x_1^{(k)},x_2^{(k)},\cdots,x_n^{(k)})^{\mathrm T},k=1,2,\cdots,$$

直接看出 $|x^{(k)}|\to 0\Leftrightarrow x_i^{(k)}\to 0(k\to\infty,1\leqslant i\leqslant n)$. 这就表明，$\mathbf{K}^n$ 中的范数收敛原不过是通常的**依坐标收敛**. 其次，对于序列 $\{u_n\}\subset B(\Omega)$，有

$$\sup_{x\in\Omega}|u_n(x)|\to 0\Leftrightarrow u_n(x)\rightrightarrows 0\quad(n\to\infty,x\in\Omega),$$

这意味着空间 $B(\Omega)$ 中的范数收敛就是**一致收敛**.

在下节中，我们将解释更多的具体空间中的范数收敛.

读者从微积分学课程中熟识的许多极限性质，如极限的惟一性与收敛序列的有界性、极限的运算性质等，都可以推广到赋范空间中的序列极限，且无需逐一重新论证，只需指明在 Euclid 空间中所述性质的证明仅用到模长性质(i)～(iii)(对

应于范数公理$(N_1)\sim(N_3)$)就够了.试用一典型例子解释如下.设在 X 中 $x_n\to x$,在 $\mathbf{K}$ 中 $\alpha_n\to\alpha(n\to\infty)$,今证在 X 中 $\alpha_n x_n\to\alpha x$.为此,只要作如下推理:

$$\begin{aligned}\|\alpha_n x_n-\alpha x\| &\leqslant \|\alpha_n x_n-\alpha_n x\|+\|\alpha_n x-\alpha x\| &&(用(N_2))\\ &=|\alpha_n|\,\|x_n-x\|+|\alpha_n-\alpha|\,\|x\| &&(用(N_1))\\ &\leqslant \mathrm{const}\,\|x_n-x\|+|\alpha_n-\alpha|\,\|x\|\to 0(n\to\infty),\end{aligned}$$

其中用到 $\{\alpha_n\}$ 的有界性.这就推广了通常的极限运算性质:乘积的极限等于极限的乘积.

从微积分学中熟知,无穷级数概念不过是有限和运算与极限运算的综合.因此,在赋范空间中考虑无穷级数是很自然的.形式上,我们几乎可照搬微积分学中的定义:若 $x_n\in X(n=1,2,\cdots)$,$s_n=\sum_1^n x_i$,$s_n\to x(n\to\infty)$,则说无穷级数 $\sum x_n$ 收敛于 x,或说 x 是级数 $\sum x_n$ 的和,写作 $x=\sum x_n$.若 $\sum\|x_n\|$ 收敛,则说级数 $\sum x_n$ **绝对收敛**.

至此,读者或许会认为赋范空间中的极限论原来很简单:无非照搬老的极限定义与结论而已.但你不可过分乐观,我们马上就要指出一个不能简单照搬的极限定理.在经典极限论中,最重要的定理无疑是:

Cauchy 收敛原理 数列 $\{x_n\}$ 收敛的充要条件是

$$\lim_{m,n}|x_m-x_n|=0, \tag{1.1.8}$$

即

$$\forall\varepsilon>0,\exists N>0,\forall m,n\geqslant N,有\ |x_m-x_n|<\varepsilon.$$

不幸的是,以上结果竟不能推广于一般赋范空间,下面就是一个简单的反例.设 P 是区间 $J=[0,1]$ 上的多项式之全体,则 P 依 sup 范数(3)显然是一个赋范空间.对于 J 上的任一连续函数,例如 $u=e^x$,由数学分析中的 Weierstrass 定理,必有多项式序列 $\{u_n\}$ 一致收敛于 u,即 $\|u_n-u\|_0\to 0(n\to\infty)$,从而当 $m,n\to\infty$ 时有

$$\|u_m-u_n\|\leqslant\|u_m-u\|_0+\|u_n-u\|_0\to 0.$$

然而,限制在空间 P 内,序列 $\{u_n\}$ 却是不收敛的.

如果我们仍然希望 Cauchy 收敛原理在赋范空间理论中发挥作用,就必须将能够应用该原理的空间从一般赋范空间中明确划分出来.这就需要以下定义.

定义 1.1.4 若赋范空间 X 中的序列 $\{x_n\}$ 满足如下 **Cauchy 条件**:

$$\lim_{m,n}\|x_m - x_n\| = 0 \tag{1.1.9}$$

(对照式(1.1.8)),则称 $\{x_n\}$ 为 **Cauchy 列**. 若 X 中所有 Cauchy 列均收敛,则称 X 为**完备赋范空间**或 **Banach 空间**.

直接看出,收敛序列必为 Cauchy 列;而定义 1.1.4 则规定 Banach 空间中的 Cauchy 列必须是收敛序列. 因此可以说,Banach 空间正是**使 Cauchy 收敛原理成立的赋范空间**. 这就将赋范空间分成了两类:完备空间与不完备空间. 鉴于 Cauchy 收敛原理在经典分析中的重要性,不难理解,在泛函分析中完备性是重要的,尽管并不是处处必要的.

值得庆幸的是,应用上常见的赋范空间大多是完备的. 首先,Cauchy 收敛原理表明 $\mathbf{K}$ 是 Banach 空间. 进而容易推出 $\mathbf{K}^n$ 亦是 Banach 空间,而这又推出(依下面给出的定理 1.1.7)任何有限维赋范空间是 Banach 空间. 例 1.1.2 所述的空间 $B(\Omega)$ 可作为无限维 Banach 空间的第一个例子,其完备性的验证是简单的(然而也是典型的):设 $\{u_n\}\subset B(\Omega)$ 是一 Cauchy 列,则 $\forall x\in\Omega$, 数列 $\{u_n(x)\}$ 满足 Cauchy 条件,因而有极限 $u(x)$. 在不等式

$$|u_n(x) - u_m(x)| \leqslant \|u_n - u_m\|_0 \quad (\forall x\in\Omega)$$

两边令 $m\to\infty$ 取极限得

$$|u_n(x) - u(x)| \leqslant \overline{\lim_m}\|u_n - u_m\|_0 \quad (\forall x\in\Omega),$$

因而

$$\|u_n - u\|_0 \leqslant \overline{\lim_m}\|u_n - u_m\|_0;$$

$$\overline{\lim_n}\|u_n - u\|_0 \leqslant \overline{\lim_n}\,\overline{\lim_m}\|u_n - u_m\|_0 = 0^{①}, \quad (\text{用 Cauchy 条件})$$

这表明 $\|u_n - u\|_0 \to 0\,(n\to\infty)$, 且 $u = u_n + (u - u_n)\in B(\Omega)$. 因此, $B(\Omega)$ 是完备的.

作为完备性的第一个应用,我们指出:完备赋范空间 X 中的绝对收敛级数 $\sum x_n$ 必收敛. 事实上,令 $s_n = \sum_1^n x_i$, 若 $\sum\|x_n\|$ 收敛,则

$$\|s_n - s_m\| = \Big\|\sum_{m<i\leqslant n} x_i\Big\| \leqslant \sum_{m<i\leqslant n}\|x_i\| \to 0 \quad (n > m\to\infty),$$

这表明 $\{s_n\}$ 是 X 中的 Cauchy 列,因而必收敛. 以上结论的重大意义在于:对于 Banach 空间中绝对收敛级数收敛性之判定,并不需要建立一套新的判别法,只需

---

① 对任何 Cauchy 列 $\{x_n\}$ 容易验证 $\overline{\lim_m}\,\overline{\lim_n}\|x_m - x_n\| = 0$. 今后将多次用到此结论.

应用熟知的**正项级数收敛判别法**就够了. 例如, 若 X 是 Banach 空间, $\{x_n\} \subset X$, $\overline{\lim_n} \sqrt[n]{\| x_n \|} < 1$, 则由熟知的 Cauchy 判别法得出, 级数 $\sum x_n$ 收敛(且绝对收敛). 这一类的结论今后将多次运用而不再详细解释.

§1.1.3 子空间·积空间与同构

在线性代数中, 对于向量空间考虑了子空间、积空间与线性同构等概念. 现在我们将这些纯代数概念与范数结合起来考虑, 看能得出哪些新的结论.

设 $A \subset X$ 是一非空子集. 若 A 对于 X 中的线性运算封闭, 则称 A 为 X 的**子空间**[①], 此时 A 依 X 中的范数同样为赋范空间. 若进而假定 A 对于 X 中的极限运算亦封闭(这意味着若 $\{x_n\} \subset A, x_n \to x$, 则 $x \in A$), 则称 A 为 X 的**闭子空间**. 以下简单命题是常用的.

命题 1.1.5 设 A 是赋范空间 X 的子空间. 若 A 作为赋范空间完备, 则它是 X 的闭子空间; 若 X 完备而 A 为闭子空间, 则 A 亦完备. 因此, Banach 空间的子空间是 Banach 空间的充要条件是它为闭子空间.

证明留作习题 2.

命题 1.1.5 常用来判定空间的完备性. 例如, 设 $J = [a,b] (a < b)$, 显然 $C(J)$(J 上的连续函数之全体)是 $B(J)$ (依例 1.1.2)的子空间. 若 $\{u_n\} \subset C(J)$, $u \in B(J)$, $\| u_n - u \|_0 \to 0$, 则 $u_n \rightrightarrows u$ (一致收敛), 于是由熟知的分析定理有 $u \in C(J)$. 可见 $C(J)$ 是 $B(J)$ 的闭子空间, 因而是完备的. 在本书中, $C(J)$ 将是出现最频繁的 Banach 空间之一.

若 A, B 是 X 的子空间, 每个 $x \in X$ 有惟一分解

$$x = a + b, \quad a \in A, \quad b \in B, \tag{1.1.10}$$

则说 X 是 A 与 B 的**直和**, 记作 $X = A \oplus B$. 令 $P_A x = a, P_B x = b$(x, a, b 依式(1.1.10)), 则得到映射

$$P_A : X \to A \text{ 与 } P_B : X \to B,$$

称之为投影映射, 简称为**投影**; 更确切地说, P_A 与 P_B 是由直和分解 $X = A \oplus B$ 所决定的投影. 若 A, B 是 X 的闭子空间, 则称直和 $X = A \oplus B$ 为**拓扑直和**. 以上概念可推广到任意有限个子空间的直和的情况.

若 $X = A + B$ (这意味着每个 $x \in X$ 有分解式(1.1.10), 但不要求惟一), A 与 B 是 X 的子空间, 则 $X = A \oplus B$ 的充要条件是 $A \cap B = \{0\}$. 这实际上是一

---

① 这意味着: 对任给 $x, y \in A, \alpha \in \mathbf{K}$, 有 $x + y \in A, \alpha x \in A$. 在具体问题中, 判定某个空间的子集为子空间, 通常是平凡的事情.

个纯代数结论，在线性代数中已遇到过，本书中将多次用到此结论.

若 X_1, X_2 是 $\mathbf{K}$ 上的两个赋范空间，以 $X_1 \times X_2$ 记所有有序对 $(x_1, x_2)(x_1 \in X_1, x_2 \in X_2)$ 之全体，则 $X_1 \times X_2$ 依线性运算

$$\alpha(x_1, x_2) + \beta(y_1, y_2) = (\alpha x_1 + \beta y_1, \alpha x_2 + \beta y_2)$$

(其中 $x_i, y_i \in X_i, i = 1,2, \alpha, \beta \in \mathbf{K}$)是 $\mathbf{K}$ 上的向量空间，称为 X_1 与 X_2 的**积空间**. $X_1 \times X_2$ 中可依多种方式定义范数. 例如可令

$$\|(x_1, x_2)\| = \|x_1\| + \|x_2\|. \tag{1.1.11}$$

容易验证，如此定义的范数的确满足范数公理，且

$$\|(x_1^{(n)}, x_2^{(n)})\| \to 0 \Leftrightarrow x_i^{(n)} \to 0 \quad (n \to \infty, i = 1,2).$$

这就表明，$X_1 \times X_2$ 中的范数收敛等价于**依坐标收敛**①. 无论 $X_1 \times X_2$ 中的范数是否依(1.1.11)式定义，只要它导出的收敛是依坐标收敛，就称该范数为**积范数**. 未加说明时，在积空间中总使用积范数②.

对于任意有限个赋范空间的积空间，可类似处理. 约定

$$\prod_{i=1}^{n} X_i = X_1 \times X_2 \times \cdots \times X_n.$$

在泛函分析中，经常不免同时涉及多个赋范空间，因而发生不同空间是否可互相替换的问题. 在线性代数中，两个互相线性同构的向量空间可以认为实质上相同而无需区别. 对于两个赋范空间，则还得加上对范数的比较，这引出以下定义.

定义 1.1.6　设 X, Y 是 $\mathbf{K}$ 上的赋范空间. 若存在线性同构③ $T: X \to Y$ 与正常数 α, β，使得

$$\alpha\|x\| \leqslant \|Tx\| \leqslant \beta\|x\|^{④} \quad (\forall x \in X) \tag{1.1.12}$$

则说 X 与 Y **拓扑同构**，并称 T 为从 X 到 Y 的一个拓扑同构；若条件(1.1.12)代之以 $\|Tx\| = \|x\| (\forall x \in X)$，则称 T 为**等距同构**. 若 $T: X \to Y$ 是等距同构而 Y 是某个赋范空间 Z 的子空间，则称 $T: X \to Z$ 为**等距嵌入**；当这样的等距嵌

---

① 若 $x = (x_1, x_2) \in X_1 \times X_2$，则称 x_1, x_2 为 x 的坐标.

② $X_1 \times X_2$ 中的另一个积范数是 $\|(x_1, x_2)\| = \max\{\|x_1\|, \|x_2\|\}$，它的作用与方便性与积范数(1.1.11)式并无差别. 其他形式的积范数尚多. 不过，在理论分析中不必关心所用的是哪个积函数.

③ 即保持线性运算的双射. 这是一个纯代数概念，你可以在线性代数中找到它.

④ 对于 X 与 Y 中的范数用了同一记号 $\|\cdot\|$，并不会引起混淆：由上下文我们能判断 $x \in X, Tx \in Y$. 今后在类似情况下，也以同一记号记不同空间中的范数.

入存在时,说 X 可**等距嵌入** Z.

以 I(或 I_X)记 X 上的**单位算子**,即 $Ix = x(\forall x \in X)$. 若 $\|\cdot\|$ 与 $\|\cdot\|'$ 是 X 上的两个范数,而

$$I:(X,\|\cdot\|)\to(X,\|\cdot\|')$$

为拓扑同构,依定义 1.1.6 这意味着存在 $\alpha,\beta>0$,使得(对照式(1.1.12))

$$\alpha\|x\|\leqslant\|x\|'\leqslant\beta\|x\|,\quad(\forall x\in X)$$

则说 $\|\cdot\|$ 与 $\|\cdot\|'$ 是 X 上的**等价范数**.

由式(1.1.12)直接推出:$\|x_n\|\to0\Leftrightarrow\|Tx_n\|\to0(n\to\infty)$. 因此,空间 X 中基于极限的所有结论通过拓扑同构 T 自然地过渡到空间 Y;或者说,就涉及极限的性质而言,X 与 Y 实质上毫无区别,因而完全可互相替换.相应地,同一空间上互相等价的范数定义出同样的极限运算,因而对于所有基于极限的问题无需区别.至于互相等距同构的赋范空间,则在涉及距离的问题上亦无区别,因而可看作完全相同的赋范空间.

互相拓扑同构的赋范空间表面上可能相差很远,实际判定两个赋范空间是否拓扑同构往往不易.不过,对于有限维空间来说事情倒很简单,因我们有以下结果.

定理 1.1.7 $\mathbf{K}$ 上所有 $n(\in\mathbf{N})$ 维赋范空间拓扑同构于 Euclid 空间 $\mathbf{K}^n$,因而互相拓扑同构.

欲知定理证明的读者可参看文献[10]的§1.9.这个出色定理的意义是显而易见的.既然 $\mathbf{K}$ 上任何 n 维赋范空间拓扑同构于 $\mathbf{K}^n$,我们只需考虑 $\mathbf{K}^n$ 这个最简单的标本就够了.因 $\mathbf{K}^n$ 完备,故所有有限维赋范空间都是完备的;这又推出任何赋范空间的有限维子空间必是闭的(用命题 1.1.5).其次,将定理 1.1.7 用于 $\mathbf{K}^n$ 得出,$\mathbf{K}^n$ 中任何范数必定互相等价.例如,$\mathbf{R}^n$ 中的范数

$$|x|_1=\sum_1^n|x_i|\quad(x=(x_i)\in\mathbf{R}^n)$$

必定等价于 Euclid 范数(1),因而依此范数的收敛就是依坐标收敛.这些事实都是基本而常用的.而且,对于定理 1.1.7 的应用还不限于上述结论.一般的规则是:若空间 $\mathbf{K}^n$ 中的某个结论仅用到线性运算与极限运算,则该结论亦适用于任何有限维赋范空间!例如,$\mathbf{K}^n$ 上的线性函数是连续的,因此任何有限维赋范空间上的线性函数亦必连续.

设 X 与 Y 分别为赋范空间与 Banach 空间.若 X 等距同构于 Y 的某个子空间 Y_0,且每个 $y\in Y$ 是 Y_0 中某序列的极限(即 Y_0 在 Y 中稠密,参看后面的定义 1.3.7),则称 Y 为 X 的**完备化**.

定理 1.1.8 任何赋范空间必存在完备化;一个赋范空间的所有完备化互相

等距同构,因而本质上是惟一的.

这一重大结果的直接证明颇为繁琐,不拟给出.在§2.4.4中我们将给出一个极简短的间接证明.

§1.2　函数空间

无论为更好地理解抽象的Banach空间概念,还是从抽象空间理论的实际应用考虑,都必须尽可能熟悉一些具体空间的例子.本节就是要专门考虑一些这样的例子.在泛函分析的应用中常见的Banach空间,大多以函数空间(包括序列空间)的形式出现.其中最基本而又常用的是本节所述的两类空间:L^p空间与C^m空间,它们分属"坏函数"的空间与"好函数"的空间,二者在实际应用及其他函数空间的研究中均起基本作用.

以下设$J=[a,b](a<b)$,$1\leqslant p=q/(q-1)\leqslant\infty$(注意$p^{-1}+q^{-1}=1$,$p=1\Leftrightarrow q=\infty$,$p=\infty\Leftrightarrow q=1$),$m\in\mathbf{Z}_+$.我们首先考虑$J$上的函数空间,这种特殊情况形式上较简单,而且实际上包含了更一般情况的本质特征.

§1.2.1　空间 $L^p(J)$

对J上任一Lebesgue可测函数$u(x)$,定义其L^p**范数**为

$$\|u\|_p=\begin{cases}\left(\int_a^b|u(x)|^p\mathrm{d}x\right)^{1/p}, & 1\leqslant p<\infty,\\ \inf\limits_{A\subset J,mA=0}\sup\limits_{x\in J\setminus A}|u(x)|, & p=\infty.\end{cases}\tag{1.2.1}$$

令

$$L^p(J)=\{u: u\text{ 在}J\text{ 上可测且 }\|u\|_p<\infty\}.\tag{1.2.2}$$

当然,(1.2.1)式中的积分应理解为Lebesgue积分,而mA则是A的Lebesgue测度①.称$u\in L^p(J)(p<\infty)$为p**次可积函数**;而称$u\in L^\infty(J)$为**本性有界函数**.若$\{u_n\}\subset L^p(J)$,$u\in L^p(J)$,$\|u_n-u\|_p\to 0(n\to\infty)$,则说序列$\{u_n\}$ L^p**收敛于**u,记作$u_n\xrightarrow{L^p}u(n\to\infty)$.当$p<\infty$时,称$L^p$收敛为$p$**次平均收敛**或$p$**次平均逼近**;1次平均收敛就简称为**平均收敛**;2次平均收敛也称为**均方收敛**.当$\|u-v\|_p$充分小时,不妨说u是v的一个L^p**逼近**或L^p**近似**.这一类的术语既方

① 不熟悉Lebesgue测度与Lebesgue积分的读者,对于理解L^p空间会感到为难.实际上,我们很快要说明,$C(J)$在$L^p(J)(p<\infty)$中稠密,因而$L^p(J)(p<\infty)$可看作空间$(C(J),\|\cdot\|_p)$的完备化.这一理解可不直接用到Lebesgue积分.

便而又颇具形象性，适当地运用并不失严格性. 直观上，$u_n \xrightarrow{L^1} u(n \to \infty)$ 意味着曲线 $u = u_n(x)$ 与 $u = u(x)$ 所夹面积趋于零. 注意这不排除对某些 x，$|u_n(x) - u(x)|$ 保持很大的值(见图 1-2). 这就是说，L^1 逼近要求整体地或平均地逼近，但不保证每个局部都无例外地逼近. 可见，L^1 收敛与通常的"处处收敛"有很大差别. 当 $1 < p < \infty$ 时，L^p 收敛有类似的直观意义.

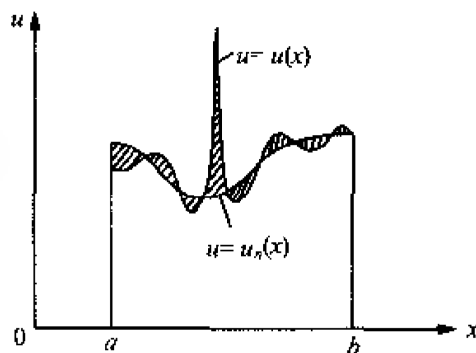


图 1-2

命题 1.2.1 若认定 J 上两个几乎处处相等的函数为同一元素，则 $L^p(J)$ 依 L^p 范数 $\|\cdot\|_p$ 是一 Banach 空间.

证 不妨设 $1 < p < \infty$($p = 1,\infty$ 的情况证明更简单). 利用不等式

$$|u + v|^p \leqslant 2^p(|u|^p + |v|^p),$$

容易验证 $L^p(J)$ 是一向量空间. 下面只需验证范数公理(N_1)~(N_3)及完备性条件(定义 1.1.4)满足.

(i) 验证范数公理. 直接看出 $\|\cdot\|_p$ 满足(N_1)与(N_3)，故只需验证(N_2). 为此，首先建立一个不等式：

$$x^\alpha y^{1-\alpha} \leqslant \alpha x + (1 - \alpha)y. \quad (x, y \geqslant 0, 0 < \alpha < 1). \tag{1.2.3}$$

不妨设 $x > y$. 固定 $y \geqslant 0$，令 $\varphi(x) = \alpha x + (1 - \alpha)y - x^\alpha y^{1-\alpha}$，则

$$\varphi'(x) = \alpha[1 - (y/x)^{1-\alpha}] > 0, \quad (x > y).$$

于是 $\varphi(x) \geqslant \varphi(y) = 0$，这推出不等式(1.2.3).

现在利用不等式(1.2.3)来建立著名的 **Hölder 不等式**：任给 $u \in L^p(J)$，$v \in L^q(J)$，有

$$\|uv\|_1 \leqslant \|u\|_p \|v\|_q. \tag{1.2.4}$$

不妨设 $\|u\|_p \|v\|_q > 0$ (否则不等式(1.2.4)自动成立)，于是

$$\begin{aligned}
\frac{\|uv\|_1}{\|u\|_p \|v\|_q} &= \int_a^b \frac{|u(x)v(x)|}{\|u\|_p \|v\|_q} dx \\
&= \int_a^b \left(\frac{|u|^p}{\|u\|_p^p}\right)^{1/p} \left(\frac{|v|^q}{\|v\|_q^q}\right)^{1/q} dx \\
&\leqslant \int_a^b \left(\frac{|u|^p}{p\|u\|_p^p} + \frac{|v|^q}{q\|v\|_q^q}\right) dx \quad (\text{取 } \alpha = p^{-1} \text{ 用式(1.2.3)})
\end{aligned}$$

$$\frac{1}{p} + \frac{1}{q} = 1,$$

这表明不等式(1.2.4)成立.

利用 Hölder 不等式,即可推出三角不等式:任给 $u, v \in L^p(J)$,有

$$\|u + v\|_p^p = \int_a^b |u + v||u + v|^{p-1}\mathrm{d}x$$

$$\leqslant \| |u||u + v|^{p-1}\|_1 + \| |v||u + v|^{p-1}\|_1$$

$$\leqslant (\|u\|_p + \|v\|_p)\| |u + v|^{p-1}\|_q \quad (\text{用式}(1.2.4))$$

$$= (\|u\|_p + \|v\|_p)\|u + v\|_p^{p/q}, \quad (\text{用}(p-1)q = p)$$

以上不等式两端同时除以 $\|u + v\|_p^{p/q}$ 即得 $\|u + v\|_p \leqslant \|u\|_p + \|v\|_p$.

(ii) 证完备性. 设 $\{u_n\} \subset L^p(J)$ 为 Cauchy 列,即

$$\lim_{m,n}\|u_m - u_n\|_p = 0.$$

则可递次取出下标 $n_1 < n_2 < \cdots$,使得

$$\|u_n - u_{n_k}\|_p < 2^{-k}. \quad (n \geqslant n_k, k = 1,2,\cdots) \tag{1.2.5}$$

令 $v_k = u_{n_{k+1}} - u_{n_k}$, $v = \sum_1^\infty |v_k|$,则

$$\|v\|_p = \left[\int_a^b \lim_n\left(\sum_1^n |v_k|\right)^p\mathrm{d}x\right]^{1/p}$$

$$= \lim_n\left[\int_a^b\left(\sum_1^n |v_k|\right)^p\mathrm{d}x\right]^{1/p} \quad (\text{用 Levi 定理})$$

$$= \lim_n\left\|\sum_1^n |v_k|\right\|_p$$

$$\leqslant \lim_n\sum_1^n\|v_k\|_p \quad (\text{用式}(1.1.4))$$

$$\leqslant \lim_n\sum_1^n 2^{-k} < \infty, \quad (\text{用式}(1.2.5))$$

可见 $v \in L^p(J)$,因而 v 在J上几乎处处有限,即 $\sum|v_k| < \infty$, a.e.. 这表明级数 $\sum v_k$ 几乎处处绝对收敛,从而序列 $\{u_{n_k}\}$ 几乎处处收敛. 设 $u_{n_k} \to u$, a.e. $(k \to \infty)$,则

$$\|u_n - u\|_p^p = \int_a^b \lim_k |u_n - u_{n_k}|^p\mathrm{d}x$$

$$\leqslant \underline{\lim_{k}}\int_a^b |u_n - u_{n_k}|^p \mathrm{d}x, \qquad \text{(用 Fatou 定理)}$$

故

$$\|u_n - u\|_p \leqslant \overline{\lim_{k}} \|u_n - u_{n_k}\|_p,$$

$$\overline{\lim_{n}} \|u_n - u\|_p \leqslant \overline{\lim_{n}}\,\overline{\lim_{k}} \|u_n - u_{n_k}\|_p = 0.$$

这表明 $\|u_n - u\|_p \to 0 (n \to \infty), u \in L^p(J)$. 完备性得证. □

以上证明的一个副产品是 Hölder 不等式(1.2.4),其应用如同 L^p 空间一样广泛,因而值得作点说明. 首先指出,不等式(1.2.4)的详细形式是

$$\int_a^b |u(x)v(x)| \mathrm{d}x \leqslant \left(\int_a^b |u(x)|^p \mathrm{d}x\right)^{1/p} \left(\int_a^b |v(x)|^q \mathrm{d}x\right)^{1/q}. \quad (1.2.4)'$$

其中 $1 < p = q/(q-1) < \infty$. 尽管式(1.2.4)′比式(1.2.4)更直观,但在熟练的情况下,运用缩写形式式(1.2.4)实际上更加方便. 下面用一个简单例子说明. 设 $1 \leqslant r < s < \infty, u \in L^s(J)$. 取 $p = s/r, q = p/(p-1), v(x) \equiv 1$, 则

$$\|u\|_r^r = \| |u|^r \|_1 = \| |u|^r v \|_1$$

$$\leqslant \| |u|^r \|_p \|v\|_q = (b-a)^{1/q} \|u\|_s^r < \infty, \quad \text{(用式(1.2.4))}$$

因而 $u \in L^r(J)$. 这就得出 $L^s(J) \subset L^r(J) (1 \leqslant r < s < \infty)$; 这一包含显然亦适用于 $s = \infty$. 特别,有 $L^p(J) \subset L^1(J) (1 < p \leqslant \infty)$. 运用 Hölder 不等式时,经常要用到以下简单事实:

$$\|u\|_p = \| |u|^p \|_1^{1/p} \quad (1 \leqslant p < \infty).$$

§1.2.2 空间 $C^m(J)$

以 $u^{(k)}$ 记函数 u 的 k 阶导数, $u^{(0)} = u$. 令

$$C^m(J) = \{u : u^{(k)} \in C(J) (0 \leqslant k \leqslant m)\},$$

约定 $C^0(J) = C(J)$. 定义

$$\|u\|_m^{①} = \max_{0 \leqslant k \leqslant m} \|u^{(k)}\|_0 \quad (u \in C^m(J)), \qquad (1.2.6)$$

其中 $\|\cdot\|_0$ 记 sup 范数(见式(1.1.3)). 对于 $u, u_n \in C^m(J) (n = 1, 2, \cdots)$, 显然

① 当 $m \geqslant 1$ 时,记号 $\|\cdot\|_m$ 可能与 L^m 范数相混. 不过,本书中 $\|\cdot\|_m$ 出现不多,且每次有明确交待,并无误解之忧.

$$\| u_n - u \|_m \to 0 \Leftrightarrow \| u_n^{(k)} - u^{(k)} \|_0 \to 0 \quad (n \to \infty, 0 \leqslant k \leqslant m)$$
$$\Leftrightarrow u_n^{(k)} \rightrightarrows u^{(k)} \quad (n \to \infty, 0 \leqslant k \leqslant m).$$

当 $\| u_n - u \|_m \to 0$ 时,我们说 $\{u_n\}$ **C^m 收敛**于 u. 这一名称并不通用,为行文方便,姑且用之.若 $\| u - v \|_m$ 充分小,则称 v 为 u 的一个 C^m 逼近或 **C^m 近似**.

对应于命题 1.2.1,现在有下面的命题.

命题 1.2.2 $C^m(J)$ 依范数 $\| \cdot \|_m$ 是一 Banach 空间.

证 $C^m(J)$ 显然是向量空间,验证 $\| \cdot \|_m$ 满足范数公理(N_1)~(N_3)也是容易的,只需证完备性.设 $\{u_n\}$ 是 $C^m(J)$ 中的 Cauchy 列,即

$$\lim_{n,l} \| u_n - u_l \|_m = 0,$$

这相当于

$$\lim_{n,l} \| u_n^{(k)} - u_l^{(k)} \|_0 = 0 \quad (0 \leqslant k \leqslant m),$$

即 $\{u_n^{(k)} : n \in \mathbf{N}\}(0 \leqslant k \leqslant m)$ 依 sup 范数为 Cauchy 列.上节中已指明 $(C(J), \| \cdot \|_0)$完备,因此有 $v_k \in C(J)$,使得 $u_n^{(k)} \rightrightarrows v_k(n \to \infty, 0 \leqslant k \leqslant m)$. 结合微积分学中的结果推出 $v_k = v'_{k-1}(1 \leqslant k \leqslant m)$. 令 $u = v_0$,则得 $v_k = u^{(k)}$, $u_n^{(k)} \rightrightarrows u^{(k)}(n \to \infty, 0 \leqslant k \leqslant m)$. 这表明 $u \in C^m(J)$ 且 $\| u_n - u \|_m \to 0(n \to \infty)$. 完备性得证. □

将已考虑的两族空间 $L^p(J)$ 与 $C^m(J)$ 合在一起,构成一个递升的空间链:

$$\cdots C^m(J) \subset C^{m-1}(J) \subset \cdots \subset C(J)$$
$$\subset L^\infty(J) \subset L^p(J) \subset L^r(J) \subset L^1(J), \tag{1.2.7}$$

其中 $1 < r < p < \infty$. 在式(1.2.7)中任取一个空间 X,例如 $X = C^2(J)$;然后任取序列 $\{u_n\} \subset X$,则

$$\{u_n\} \subset C^1(J) \subset C(J) \subset L^p(J) \quad (1 \leqslant p \leqslant \infty).$$

因此,对于 $\{u_n\}$ 可以考虑 C^m 收敛 $(m = 2,1,0)$ 与 L^p 收敛 $(1 \leqslant p \leqslant \infty)$,这就得到无限多种收敛性;而且,这些收敛性的强度是递降的.

$$C^2 \text{ 收敛} \Rightarrow C^1 \text{ 收敛} \to C^0 \text{ 收敛} \Rightarrow L^p \text{ 收敛} \Rightarrow L^r \text{ 收敛},$$

其中 $1 \leqslant r < p \leqslant \infty$ [①]. 可能使你困惑的是:为何要这许多收敛性? 简单的回答

① 显然 L^∞ 收敛⇒L^r 收敛;若 $1 \leqslant r < p < \infty$,则由 Hölder 不等式有 $\| u \|_r^r \leqslant \text{const} \| u \|_p^r$,由此易看出 L^p 收敛⇒L^r 收敛.

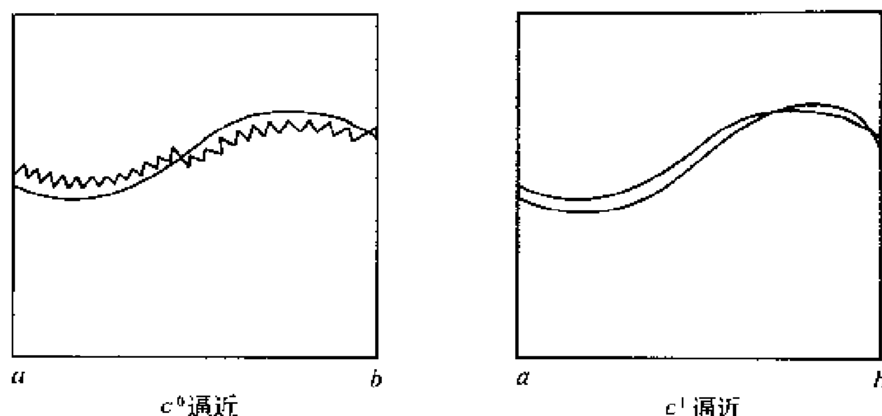


图 1-3

是:惟此才足以应付多种不同的需要. 例如,设函数 $u(\cdot)$ 描述了某个宏观的物理量,当你并不关心局部问题时,$u_n \xrightarrow{L^1} u$ 已足以使你满意. 若 $u \in C^1(J)$,且你除了关心 $u(\cdot)$ 的数值之外还关心曲线 $u = u(x)$ 的光滑性,那么仅仅 $u_n \rightrightarrows u$ 并不能使你满意,你会同时要求 $u'_n \rightrightarrows u'$, 这就需要 C^1 收敛了如图 1-3. 对于更精细的问题,需要 C^2 收敛甚至更强的收敛.

§1.2.3　各种推广

对于应用上需要的函数空间,当然不能将函数的定义域限制为区间 J. 现在就解除这一限制,考虑更一般的空间 $L^p(\Omega)$ 与 $C^m(\Omega)$. 不过,这种推广并不涉及本质上新的因素,只是形式上更繁些罢了. 因此,我们将省去大部分细节,仅概述基本的定义与结论.

(i) **空间 $L^p(\Omega)$**. 设 (Ω, μ) 是一测度空间. 对 Ω 上任一可测函数 u, 令

$$\| u \|_p = \begin{cases} \left(\int_\Omega | u |^p \mathrm{d}\mu\right)^{1/p}, & 1 \leqslant p < \infty, \\ \inf\limits_{A \subset \Omega, \mu A = 0} \sup\limits_{x \in \Omega \setminus A} | u(x) |, & p = \infty, \end{cases} \tag{1.2.1$'$}$$

$$L^p(\Omega) = \{u : u \text{ 是 } \Omega \text{ 上的可测函数且 } \| u \|_p < \infty\}. \tag{1.2.2$'$}$$

若 Ω 上几乎处处相等的函数不加区别,则 $(L^p(\Omega), \|\cdot\|_p)$ 是一 Banach 空间,将其记为 $L^p(\Omega, \mu)$, 简写作 $L^p(\Omega)$ 或 L^p. 相应地,前面已提到的 L^p 范数、L^p 收敛等术语在此有自明的意义,Hölder 不等式(1.2.4)亦保持成立. 若 $\Omega \subset \mathbf{R}^n$ 而未特别说明,则 $L^p(\Omega)$ 总理解为 $L^p(\Omega, m)$,m 为 n 维 Lebesgue 测度. 若 Ω 是一个概率空间, P 是其中的概率测度,则 $L^p(\Omega)(= L^p(\Omega, P))$ 就是 Ω 上 p 阶随机变量

之全体，而 $\|u\|_p = (E|u|^p)^{1/p}$，$E|u|^p$ 是随机变量 u 的 p 阶绝对矩 $(1 \leqslant p < \infty)$，$u \in L^p(\Omega)$ 也称为 p 次可积随机变量. 实际上，关于随机变量以及随机过程的相当一部分内容，正是在 L^p 空间（尤其是在 L^2 空间）的框架内得到有效处理.

(ii) **空间 l^p**. 任给无穷向量 $x = (x_1, x_2, \cdots, x_i, \cdots)^T \triangleq (x_i)$，令

$$\|x\|_p = \begin{cases} \left(\sum_i |x_i|^p\right)^{1/p}, & 1 \leqslant p < \infty, \\ \sup_i |x_i|, & p = \infty \end{cases} \tag{1.2.8}$$

则

$$l^p = \{x = (x_i) : \|x\|_p < \infty\} \quad (1 \leqslant p \leqslant \infty) \tag{1.2.9}$$

依范数 $\|\cdot\|_p$ 是一个 Banach 空间. 空间 l^∞ 已在 §1.1.1 中提到了. l^p 是最简单、因而最先被研究的无限维 Banach 空间之一，常用作解释某些概念与结论的模型. 在很多方面，可将空间 l^p 与 L^p 相对照. 例如，与式(1.2.4)相对应，成立如下形式的 Hölder 不等式：

$$\left|\sum_i x_i y_i\right| \leqslant \|x\|_p \|y\|_q, \tag{1.2.10}$$

其中 $x = (x_i) \in l^p$，$y = (y_i) \in l^q$，$1 \leqslant p = q/(q-1) \leqslant \infty$. 特别，取 $p = q = 2$ 得到 **Cauchy 不等式**：

$$\left|\sum_i x_i y_i\right|^2 \leqslant \sum_i |x_i|^2 \sum_i |y_i|^2, \tag{1.2.11}$$

其中 $(x_i), (y_i) \in l^2$.

(iii) **空间 $C^m(\Omega)$**. 设 $\Omega \subset \mathbf{R}^n$ 是一有界闭区域，以 $C^m(\Omega)$ 记 Ω 上 m 次连续可微（即所有 m 阶偏导数存在且连续）函数之全体，令

$$\|u\|_m = \max_{|\alpha| \leqslant m} \|\partial^\alpha u\|_0 \quad (u \in C^m(\Omega)), \tag{1.2.6'}$$

其中

$$\begin{cases} \partial^\alpha u = \dfrac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_n^{\alpha_n}}, \\ \alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n) \in \mathbf{Z}_+^n, \ |\alpha| = \sum_i \alpha_i. \end{cases} \tag{1.2.12}$$

自然约定 $\partial^0 u = u$，此处 $0 = (0, 0, \cdots, 0) \in \mathbf{Z}^n$. 可以验证 $(C^m(\Omega), \|\cdot\|_m)$ 是一 Banach 空间，在其中 $u_k \to u\ (k \to \infty)$ 意味着在 Ω 上

$$\partial^\alpha u_k \rightrightarrows \partial^\alpha u \quad (k \to \infty, |\alpha| \leqslant m).$$

如在 $C^m(J)$ 中一样，也称上述收敛为 C^m 收敛.

若空间 $L^p(\Omega)$，$C^m(\Omega)$ 中的函数取实(或复)值，则 $L^p(\Omega)$，$C^m(\Omega)$ 是实(或复)Banach 空间. 对于空间 l^p 亦可作类似说明. 不过，涉及这些空间的许多结论与所取数域无关.

最后应当指出，应用函数空间来描述具体的物理或技术系统(甚至包括某些社会系统)，实际上是很自然的. 例如，一个物理系统可能具有的状态通常决定于若干参数(时间、空间位置及其他因素)，特定的状态一般对应某一特定的函数(或序列)，函数的类型(可积、连续或光滑等)则决定于系统的具体特性. 如果我们希望整体地把握一个系统，就不应仅对某些特殊的状态感兴趣，而应对系统可能有的状态之全体感兴趣，而这就涉及到一定的函数空间. 实际上，对任何物理系统的考察都必然地关联着相应的函数空间，只是有时未必明确意识到罢了.

§1.3　点集与连续性

本节的任务是：借助于平常的几何术语，对赋范空间中点集的构造给出某种形象化的描述；进而运用点集论的概念、语言与方法，处理与连续性及与空间结构有关的问题. 对于本节，你的最初印象很可能是：罗列的概念很多，而重大结论则较少. 然而，所述内容对于整个泛函分析都是很基本的，今后几乎随时用到.

以下设 X，Y 等是给定的赋范空间.

§1.3.1　点集

空间 X 的任何子集称为**点集**，其中的元素称为**点**. 你自然想像点集如同平常空间(例如平面 $\mathbf{R}^2$)中的图形. 在 $\mathbf{R}^2$ 中，我们可说出三角形、椭圆等许多熟知的图形. 然而，对于抽象的空间 X，似乎还说不出其中有什么特别令人感兴趣的点集. 不过，有一种点集，即球，是特别简单且容易想像的. 任给 $a \in X$、$r > 0$，令

$$B_r(a) = \{x \in X : \| x - a \| < r\}, \tag{1.3.1}$$

$$\bar{B}_r(a) = \{x \in X : \| x - a \| \leqslant r\}. \tag{1.3.2}$$

二者分别称为 X 中以 a 为心以 r 为半径的**开球**与**闭球**，开球也简称为**球**；称 $B_1(0)(\bar{B}_1(0))$ 为**单位球**(闭单位球). 有时也用到**球面**概念，并约定 $S_r(a) = \{x \in X : \| x - a \| = r\}$. 我们将看到，球是赋范空间理论中一个特别有用的辅助工具.

通过类比与联想，你很容易想像，赋范空间中的球就是与通常的球或 $\mathbf{R}^2$ 中的圆相类似的东西. 这一比拟大致不错，但也不可过于简单化，以免发生误导. 即使在有限维空间中，球概念也比初看起来要复杂些. 例如，在 $\mathbf{R}^2$ 中定义

$$\begin{cases} \| x \|_p = (| x_1 |^p + | x_2 |^p)^{1/p} & (1 \leqslant p < \infty), \\ \| x \|_\infty = \max\{| x_1 |, | x_2 |\}, \end{cases}$$

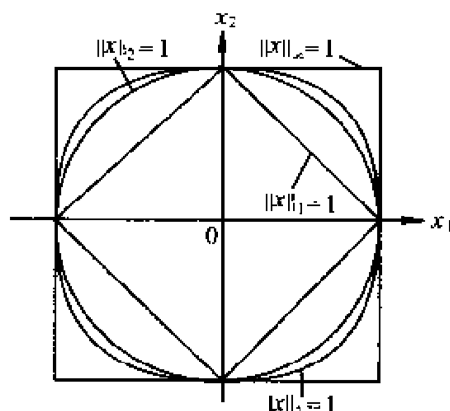


图 1-4

$(x=(x_1,x_2)\in\mathbf{R}^2)$，则对于不同的 p，单位球面 $\{x\in\mathbf{R}^2:\|x\|_p=1\}$ 是很不相同的(见图 1-4)．顺便指出：正是球的不同特性，决定了空间 X 的不同类型．

如果 $A\subset X$ 是一个任意的点集，那么能对它说些什么呢？我们能将从观察通常的平面图形所获得的印象加于 A 吗？下面的定义正反映了我们力图类比于平常的几何观察来描述抽象点集的意向．

定义 1.3.1　设 $A\subset X$，$x\in X$．

(i) 若存在 $r>0$，使得 $B_r(x)\subset A$，则称 x 为 A 的**内点**，称 A 为 x 的**邻域**①．以 A° 记 A 的内点之全体，称它为 A 的**内部**．

(ii) 若 $\forall r>0$，有 $A\cap B_r(x)\neq\varnothing$，则称 x 为 A 的**触点**．以 $\bar{A}$ 记 A 的触点之全体，称它为 A 的**闭包**．

(iii) 若 $\forall r>0$，有 $A\cap B_r(x)\neq\varnothing$，$A^c\cap B_r(x)\neq\varnothing$，则称 x 为 A 的**边界点**．以 ∂A 记 A 的边界点之全体，称它为 A 的**边界**．

(iv) 若 $\forall r>0$，有 $(A\setminus\{x\})\cap B_r(x)\neq\varnothing$，则称 x 为 A 的**聚点**或**极限点**．以 A' 记 A 的聚点之全体，称它为 A 的**导集**．

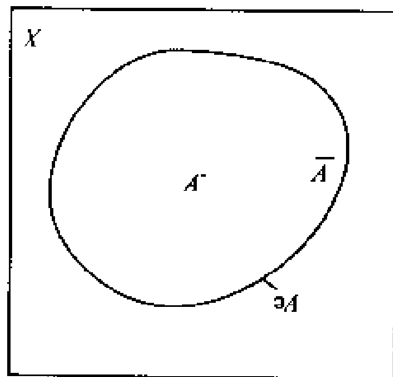


图 1-5

这样，每个点集 A 都关联着四个集：$A^\circ,\bar{A},\partial A$ 与 A'．将集 $A^\circ,\partial A$ 及 $\bar{A}=A^\circ\cup\partial A$ 想像成如图1-5所示的形状，或许过于简单化了．然而，正是从如图 1-5 这样简单的示意图所获得的直观印象，为理解抽象的点集概念提供了有益启示．只要你不过分依赖于直观，就不致陷入谬误．另一方面，利用定义 1.3.1 可推出以下关系：

$$
\begin{cases}
\bar{A}=A^{\text{cx}}=A^\circ\cup\partial A=A\cup A', & (1.3.3\text{a})\\
A^\circ=(\overline{A^c})^c=\bar{A}\setminus\partial A=A\setminus\partial A, & (1.3.3\text{b})\\
\partial A=\bar{A}\setminus A^\circ=\bar{A}\cap\overline{A^c}, & (1.3.3\text{c})\\
A'=\{x:x\in\overline{A\setminus\{x\}}\}. & (1.3.3\text{d})
\end{cases}
$$

① 较常用的是开邻域，特别是球形邻域．但限于使用开邻域并无必要．

这些关系式不如定义 1.3.1 中的直接描述那么直观，但用于逻辑证明则是更方便的.不过，眼下使你感到烦恼的可能是，突然一下子面对一大堆既琐碎又欠直观的公式，不知如何能够记住.其实，你未必需要去记住它们.但你不妨自己动手证明其中部分结论，以便加深理解.式(1.3.3a)～(1.3.3d)的证明都很简单，今以证 $A^\circ = (\overline{A^c})^c$ 作为示例.

$$x \in A^\circ \Leftrightarrow \exists\, r > 0 : B_r(x) \subset A \qquad (\text{用定义 } 1.3.1(\mathrm{i}))$$

$$\Leftrightarrow \exists\, r > 0 : A^c \cap B_r(x) = \varnothing$$

$$\Leftrightarrow x \notin \overline{A^c} \Leftrightarrow x \in (\overline{A^c})^c. \qquad (\text{用定义 } 1.3.1(\mathrm{ii}))$$

式(1.3.3)表明，$A^\circ, \partial A$ 与 A' 都可通过闭包来表示，可见闭包是一个基本的概念，值得对它作更细致的考察.

命题 1.3.2 设 $A, B \subset X, x \in X$，则以下结论成立：

(i) $x \in \overline{A} \Leftrightarrow d(x, A) = 0 \Leftrightarrow$ 存在序列 $\{x_n\} \subset A$，使 $x_n \to x (n \to \infty)$.

(ii) $A \subset \overline{A}; A \subset B \Rightarrow \overline{A} \subset \overline{B}; \overline{A} = \overline{\overline{A}}; \overline{A \cup B} = \overline{A} \cup \overline{B}$.

证 (i) 若 $x \in \overline{A}$，则 $\forall\, r > 0$，有 $a \in A \cap B_r(x)$，于是

$$d(x, A) \leqslant \| x - a \| < r,$$

这推出 $d(x, A) = 0$. 若 $d(x, A) = 0$，则 $\forall\, n \in \mathbf{N}$，必有 $x_n \in A$，使 $\| x_n - x \| < 1/n$，这表明 $\{x_n\} \subset A, x_n \to x$. 若存在 $\{x_n\} \subset A$ 使 $x_n \to x (n \to \infty)$，则对任给 $r > 0$，当 n 充分大时 $\| x_n - x \| < r$，因而 $x_n \in A \cap B_r(x) \neq \varnothing$，这推出 $x \in \overline{A}$.

(ii) 显然 $A \subset \overline{A}, A \subset B \Rightarrow \overline{A} \subset \overline{B}, \overline{A} \subset \overline{\overline{A}}$. 若 $x \in \overline{\overline{A}}$，则由已证的(i)有 $\{x_n\} \subset \overline{A}$，使 $x_n \to x (n \to \infty)$. 取 $y_n \in A \cap B_{1/n}(x_n) (n = 1, 2, \cdots)$，则显然 $y_n \to x (n \to \infty)$，故 $x \in \overline{A}$. 这就证得 $\overline{A} = \overline{\overline{A}}$. 由 $A \subset A \cup B$ 得 $\overline{A} \subset \overline{A \cup B}$；同理 $\overline{B} \subset \overline{A \cup B}$，因此 $\overline{A} \cup \overline{B} \subset \overline{A \cup B}$. 若 $x \in \overline{A \cup B}$，则有 $\{x_n\} \subset A \cup B$，使 $x_n \to x (n \to \infty)$. 不妨设有 $\{x_n\}$ 的子列 $\{x_{n_k}\} \subset A$，因此由 $x_{n_k} \to x$ 得 $x \in \overline{A}$，从而 $x \in \overline{A} \cup \overline{B}$，这就证得 $\overline{A \cup B} = \overline{A} \cup \overline{B}$. □

现在引进点集论中最重要的概念.

定义 1.3.3 设 $A \subset X$. 若 $A = \overline{A}$，则称 A 为闭集；若 $A = A^\circ$，则称 A 为开集.

今将闭集与开集的性质综合于下.

命题 1.3.4 设 $A \subset X$，则以下结论成立：

(i) 互补性：A 是闭集 $\Leftrightarrow A^c$ 是开集.

(ii) 等价刻画：A 是闭集 $\Leftrightarrow \overline{A} \subset A \Leftrightarrow \partial A \subset A \Leftrightarrow A' \subset A \Leftrightarrow A$ 对极限运算封

闭,即 A 中任何收敛序列的极限属于 A;A 是开集 $\Leftrightarrow A \subset A^\circ \Leftrightarrow A \cap \partial A = \varnothing$.

(iii) 运算性质:X 中任意个闭集的交或有限个闭集的并是闭集,任意个开集的并或有限个开集的交是开集.

证 (i) A 是闭集 $\Leftrightarrow A = \bar{A} = A^{cc} \Leftrightarrow A^c = A^{c\circ} \Leftrightarrow A^c$ 是开集(用式(1.3.3a)).由于开集与闭集之间的互补关系,下面只需考虑闭集.

(ii) 由式(1.3.3)及命题 1.3.2 之(i)推出(试推证).

(iii) 设 $\{A_i\}$ 是 X 中任一族闭集,$A = \bigcap_i A_i$,则由 $A \subset A_i$ 得 $\bar{A} \subset \bar{A}_i = A_i$,从而 $\bar{A} \subset \bigcap_i A_i = A$,故 A 是闭集.其次设 $A, B \subset X$ 是闭集,则由(利用命题1.3.2之(ii))

$$\overline{A \cup B} = \bar{A} \cup \bar{B} = A \cup B$$

知 $A \cup B$ 是闭集.进而用归纳法可知任意有限个闭集之并为闭集. □

简单地说,闭集就是对极限运算封闭的集;开集就是其中每点为内点的集.你对于闭集的印象很可能是如同用边界围起来的闭区域那样的集.这种理解并不准确,容易导致误解.一个闭集可以没有边界(即边界为空集),例如全空间 X 就是,它并不被什么边界围住!一个闭集可能其本身就是边界,例如若 A 是 $\mathbf{R}^2$ 中一直线,则 $A = \partial A$!没有聚点的集必然是闭集($A' = \varnothing \Rightarrow A' \subset A$!),例如 $\mathbf{R}$ 的子集 $\mathbf{N}$ 就是闭集,它看起来更不像闭区域或闭区间.

在具体问题中,已知某个集是闭集或开集常导致极有用的定性判断.例如,在空间 $C(J)$ 中,集

$$A \triangleq \{u \in C(J): u \text{ 在 } J \text{ 上有零点}\}$$

是闭集(请证之!).因此,若 $\{u_n\} \subset A$,$u_n \to u$,则必 $u \in A$,即 u 在 J 上必有零点.再如,在 $n \times n$ 阶实矩阵的空间 $\mathbf{R}^{n \times n}$ 中,可逆矩阵之全体构成一开集 G.因此,若 $M \in G$,则充分接近于 M 的矩阵也属于 G;或者说,微小的扰动不会使 M 逸出集 G 之外.这就说明,矩阵的可逆性是一**稳定性质**,它不因微小扰动而被破坏.这一类的结论在实际问题中无疑是很有意义的.

§1.3.2 连续性

首先概述一下有关映射的基本用语.设 $D \subset X$ 是一非空集.若对每个 $x \in D$,依规则 F 指定了一点 $y \in Y$ 与之对应(通常写 y 为 $F(x)$,或简写作 Fx),则说给

定了一个映射[①] $F: D \to Y$，且称 D 为 F 的**定义域**，记作 $D(F)$，称 $R(F) \triangleq \{Fx: x \in D\}$ 为 F 的**值域**，称 Fx 为 F 在 x 的**值**；当 $Fx \in \mathbf{K}$ 时通常将 Fx 写作 $F(x)$. 当要突出 D 是 X 的子集这一事实时，将 $F: D \to Y$ 写成

$$F: D \subset X \to Y.$$

设 $F: D \to Y$ 是一映射. 若 $Fx = Fy \Rightarrow x = y$，则称 F 为**单射**；若 $R(F) = Y$，则称 F 为**满射**；同时为单射与满射的映射称为**双射**. 若 F 为双射，则 F 有**逆映射** F^{-1}，即

$$F^{-1}: Y \to D, \quad Fx \to x.$$

若 $\varnothing \neq A \subset D$，则称映射 $A \to Y, x \to Fx$ 为映射 $F: D \to Y$ 在 A **上的限制**，记作 $F \mid A$；而称 F 为 $F \mid A$ 的**延拓**或**扩张**. 若 $G: Y \to Z$ 是另一映射，则称映射 $D \to Z, x \to G(Fx)$ 为 G 与 F 的**复合映射**，记作 $G \circ F$. 对于映射 $F: D \to Y$ 与 $A \subset D, B \subset Y$，约定

$$\begin{cases} FA = \{Fx : x \in A\}, \\ F^{-1}B = \{x \in D : Fx \in B\}. \end{cases}$$

FA 称为 A 在映射 F 下的**像**，$F^{-1}B$ 称为 B 关于 F 的**原像**. 关于像与原像的以下简单关系是常用的.

$$\begin{cases} A \subset F^{-1}(FA), \quad F(F^{-1}B) \subset B, & (1.3.4a) \\ A \subset F^{-1}B \Leftrightarrow FA \subset B, & (1.3.4b) \\ F(\bigcup_i A_i) = \bigcup_i FA_i, \quad F(\bigcap_i A_i) \subset \bigcap_i FA_i, & (1.3.4c) \\ F^{-1}(\bigcup_i B_i) = \bigcup_i F^{-1}B_i, \quad F^{-1}(\bigcap_i B_i) = \bigcap_i F^{-1}B_i, & (1.3.4d) \\ F^{-1}B^c = D \setminus F^{-1}B, & (1.3.4e) \\ (G \circ F)^{-1}C = F^{-1}(G^{-1}C), & (1.3.4f) \end{cases}$$

其中 $F: D \to Y, G: Y \to Z, A, A_i \subset D, B, B_i \subset Y (i \in I, I$ 是任一非空指标集$)$，$C \subset Z$.

作了以上准备之后，现在开始考虑连续映射. 几乎照搬微积分学中的连续性定

① 在现代数学中，映射、函数、算子、变换等用语并无实质上的差别，只是因习惯的关系，在不同的场合，用法上有所偏重. 例如，泛函分析中通用算子一词；而取值于 $\mathbf{R}$ 或 $\mathbf{C}$ 中的映射通常叫函数或泛函. 名称的选择并不是本质的.

义得到：

定义 1.3.5　给定映射 $F: D\subset X\to Y$ 与点 $x_0\in D$. 若 $\forall\varepsilon>0,\exists\delta>0$, 当 $x\in D$, $\|x-x_0\|<\delta$ 时恒有 $\|Fx-Fx_0\|<\varepsilon$, 则说 F 在点 x_0 处连续. 若 F 在 D 上每点连续, 则说 F 在 D 上连续.

约定以 $C(D,Y)$ 记从 D 到 Y 的连续映射之全体, 而令 $C(D)=C(D,\mathbf{K})$, $\mathbf{K}=\mathbf{R}$ 或 $\mathbf{C}$. 对于 $f\in C(D)$, 更习惯的叫法是称 f 为连续函数或连续泛函.

运用点集记号, 可将 F 在 x_0 连续的条件表成: $\forall\varepsilon>0,\exists\delta>0$, 使

$$F(D\cap B_\delta(x_0))\subset B_\varepsilon(Fx_0)^{①}. \tag{1.3.5}$$

容易说明, 以上条件等价于: 若在 D 中 $x_n\to x_0$, 则 $Fx_n\to Fx_0(n\to\infty)$, 即

$$\lim_n Fx_n=F(\lim_n x_n). \tag{1.3.6}$$

式(1.3.6)正是我们从微积分学中熟知的表达连续性的标准方式. 这样, 我们在微积分学中所熟悉的所有连续函数, 都可以用作定义 1.3.5 意义下的连续函数的例子. 当然, 你更希望看到一些一般赋范空间 X 上的连续函数的例子. 可惜, 眼下就能写出的这种例子还不多. 一个最常用的简单例子就是 $f(x)=\|x\|$, 它的连续性是很明显的(请试证明).

在泛函分析中, 基于点集论的以下连续性条件或许更加有用.

定理 1.3.6　设 $D\subset X$ 是开(或闭)集.

(i) 映射 $F: D\to Y$ 连续的充要条件是, 对任给开(或闭)集 $V\subset Y$, $F^{-1}V$ 是 X 中的开(或闭)集.

(ii) 函数 $f: D\to\mathbf{R}$ 连续的充要条件是, $\forall\alpha\in\mathbf{R}$, 集 $D(f<\alpha)$ 与 $D(f>\alpha)$ 为开集(或 $D(f\leqslant\alpha)$ 与 $D(f\geqslant\alpha)$ 为闭集), 其中

$$D(f<\alpha)=\{x\in D: f(x)<\alpha\}=f^{-1}(-\infty,\alpha);$$

$D(f>\alpha)$ 与 $D(f\leqslant\alpha)$, $D(f\geqslant\alpha)$ 仿此.

证　不妨只考虑 D 为开集的情况(利用式(1.3.4c)可过渡到 D 为闭集的情况).

(i) 首先设 F 连续, $V\subset Y$ 是一开集. 任给 $x_0\in F^{-1}V$, 今要证 x_0 是 $F^{-1}V$ 的内点(这推出 $F^{-1}V$ 为开集). 因 $Fx_0\in V=V^\circ$, 故有 $\varepsilon>0$, 使 $B_\varepsilon(Fx_0)\subset V$ (依定义 1.3.1(i)). 由连续性有 $\delta>0$, 使条件(1.3.5)满足. 因 D 是开集, 故不妨设 $B_\delta(x_0)\subset D$ (否则适当缩小 δ), 因而由式(1.3.5)有 $F(B_\delta(x_0))\subset V$, 这推出

① 注意球 $B_\delta(x_0)$ 与 $B_\varepsilon(Fx_0)$ 分别在空间 X 与 Y 中. 此处及以后对于不同空间的球都用同一记号 $B_r(a)$, 所属空间则由球心 a 来识别, 并无混淆之虞.

$B_\delta(x_0) \subset F^{-1}V$ (用式(1.3.4b)),故 $x_0 \in (F^{-1}V)^\circ$.

反之,设开集关于 F 的原像恒为开集,今证 F 在任一点 $x_0 \in D$ 连续.任给 $\varepsilon > 0$, $V \triangleq B_\varepsilon(Fx_0)$ 是 Y 中的开集,因而 $F^{-1}V$ 是 X 中的开集,故 x_0 是 $F^{-1}V$ 的内点.于是有 $\delta > 0$,使 $B_\delta(x_0) \subset F^{-1}V$,从而 $FB_\delta(x_0) \subset V$,即条件(1.3.5)满足,可见 F 在 x_0 处连续.

(ii) 若 f 连续,则由已证的(i)知 $D(f < \alpha) = f^{-1}(-\infty, \alpha)$ 与 $D(f > \alpha) = f^{-1}(\alpha, \infty)$ 是开集.反之,设 $D(f < \alpha)$ 与 $D(f > \alpha)(\alpha \in \mathbf{R})$ 恒为开集,则对任何有限开区间 $(\alpha, \beta) \subset \mathbf{R}$,

$$f^{-1}(\alpha, \beta) = D(f < \beta) \cap D(f > \alpha)$$

是 X 中的开集.因 R 中的任何开集 V 可表为有限开区间之并,故 $f^{-1}V$ 恒为开集.由已证的(i),f 连续. □

用定理 1.3.6 所给的条件来判定连续性是否方便,此处姑且不论.可以立即指出的是,定理 1.3.6 提供了构成开集与闭集的很一般的方法.例如,前面提到可逆矩阵之集 G 是 $\mathbf{R}^{n\times n}$ 中的开集,你可能正疑惑于如何得出这一结论.从定理 1.3.6 看来,这其实很简单,

$$f: \mathbf{R}^{n\times n} \to \mathbf{R}, \quad A \to \det A$$

的连续性是不成问题的,因而

$$G = \{A : f(A) \neq 0\} = f^{-1}((-\infty, 0) \cup (0, \infty))$$

是开集.一般地,若 $f_i : X \to \mathbf{R}(1 \leqslant i \leqslant n)$ 是一组连续函数,集 $E \subset X$ 由涉及 f_i 的(有限个)不等式界定,则当仅有严格不等号 $>$,$<$(或不严格不等号 $\geqslant$,$\leqslant$)出现时,E 是开集(或闭集).例如,不加思索,你就能断定

$$\{(x, y) \in \mathbf{R}^2 : x^2 + y^2 < x < 2x^2 + y\}$$

与

$$\{(x, y) \in \mathbf{R} : 2x^2 + y^2 \leqslant y \leqslant x^2 + 1\}$$

分别为 $\mathbf{R}^2$ 中的开集与闭集,尽管你未必能立即说出这些集的准确形状.但对于集

$$\{x \in X : \| x \| < 1, \| x - a \| \geqslant 1\},$$

却不能作结论.

§1.3.3 基本集

现在我们来看点集论的概念是否有助于揭示赋范空间的结构.中心的问题是,

能否找到一个尽可能“小”的子集 $A \subset X$，使得每个 $x \in X$ 可以适当形式由 A 中的元素表示出来. 若 $\dim X = n < \infty$，则可取 X 的一个基 $\{e_1, e_2, \cdots, e_n\}$，使得每个 $x \in X$ 可表为 $e_i (1 \leqslant i \leqslant n)$ 的线性组合，即

$$X = \operatorname{span}\{e_1, e_2, \cdots, e_n\},$$

记号 $\operatorname{span}A$ 表示集 A 生成的向量空间，即 A 中元的有限线性组合之全体. 因此，在某种意义上，对于 X 的研究，仅使用 n 个元 $e_1, e_2, \cdots, e_n$ 就够了. 如此带来的简化何其明显. 若 $\dim X = \infty$，则已不可能找到由有限个元组成的基. 但我们仍然力图取出较小的子集(例如可数集) $A \subset X$，使得每个 $x \in X$ 可用 A 中的元通过线性运算与极限运算表示出来，因而某些问题只需在 A 上处理就够了. 这一考虑导致一系列概念，它们在 Banach 空间理论中具有基本的重要性.

定义 1.3.7　设 $A, B \subset X$.

(i) 若 $B \subset \bar{A}$，则说 A 在 B 中**稠密**；若 $A \subset B \subset \bar{A}$，则称 A 为 B 的稠子集；X 的稠子集就称为**稠集**.

(ii) 若 B 含可数的稠子集，则称 B 为**可分集**；若 X 本身可分，则称 X 为**可分空间**.

(iii) 若 $\overline{\operatorname{span}A} = X$，即 $\operatorname{span}A$ 为稠集，则称 A 为 X 的**基本集**.

(iv) 若 $\{e_i\} \subset X$ 是一序列，每个 $x \in X$ 可惟一地表为

$$x = \sum_i \alpha_i e_i,$$

则称 $\{e_i\}$ 为 X 的 **Schauder 基**.

直接从定义看出，若 A 是 X 中的稠集，则每个 $x \in X$ 可表为 A 中某序列的极限(参考命题 1.3.2(i))，或说 x 可用 A 中的元逼近；若 A 是 X 的基本集，则每个 $x \in X$ 可用 A 中元的线性组合逼近. 稠集与 Schauder 基都是基本集；X 可分$\Leftrightarrow$ X 有可数的基本集，因而有 Schauder 基的空间必定可分.

在定义 1.3.7 所定义的诸概念中，最重要的是基本集，它可看作与有限维空间的基相当的东西. 显然，任何赋范空间 X 都有基本集—— X 本身就是. 但如果一个基本集不足够“小”，或其中的元素没什么特色，就未必有什么价值. 我们当然要寻求更有价值的基本集. 现在就来考虑一些具体空间中基本集的例子.

例 1.3.8　(i) 空间 l^p 的基本集. 令 $e_i = (0, \cdots, 0, 1, 0, \cdots)^{\mathrm{T}}$，1 在第 i 项，今指明 $\{e_i : i \in \mathbf{N}\}$ 是空间 $l^p (1 \leqslant p < \infty)$ 的 Schauder 基，因而是 l^p 的基本集且 l^p 是可分的. 事实上，任给 $x = (x_i) \in l^p$，有

$$\left\| x - \sum_{i=1}^{n} x_i e_i \right\|_p^p = \sum_{i > n} |x_i|^p \to 0 \quad (n \to \infty),$$

这正表明 $x=\sum_1^\infty x_i e_i$.

为行文方便起见，将如上的 $\{e_i\}$ 称为 l^p 的**标准基**. 今后当提及 l^p 与 $\{e_i\}$ 而未加说明时，总认定 $\{e_i\}$ 就是标准基.

(ii) 空间 $C(J)$ 的基本集，$J=[a,b](a<b)$. 由著名的 Weierstrass 定理，每个 $u\in C(J)$ 可用 J 上的多项式一致逼近，这等于说，J 上的多项式之全体 P 是 $C(J)$ 中的稠集. 其次，以 A 记幂函数 $x^n(n\in \mathbf{Z}_+)$ 之全体，则显然 $P=\mathrm{span}A$，故 A 是 $C(J)$ 的基本集，因而 $C(J)$ 是可分的.

(iii) 空间 $L^p(J)$ 的基本集，J 如上，$1\leqslant p<\infty$. 因可指明每个 $u\in L^p(J)$ 可用连续函数 L^p 逼近，而一致逼近强于 L^p 逼近(参考 §1.2.2)，故空间 $C(J)$ 的基本集 $A=\{x^n:n\in \mathbf{Z}_+\}$ 也是 $L^p(J)$ 的基本集，因而 $L^p(J)$ 是可分的. 其次，因易指明每个 $u\in C(J)$ 可用阶梯函数一致逼近，而阶梯函数是形如 χ_δ(δ 是 J 的子区间)的函数的线性组合，故

$$\{\chi_\delta:\delta \text{ 是 } J \text{ 的子区间}\} \tag{1.3.7}$$

亦为 $L^p(J)$ 的基本集. 其次容易看出，也可以 $\{\chi_{[a,x]}:x\in J\}$ 代替式(1.3.7)作为 $L^p(J)$ 的基本集.

(iv) 空间 $L^p(\mathbf{R})$ 的基本集，$1\leqslant p<\infty$. 对任给 $u\in L^p(\mathbf{R})$，令 $u_n=u\chi_{[-n,n]}$，则

$$\|u_n-u\|_p^p=\int_{|x|>n}|u(x)|^p\mathrm{d}x\to 0\quad(n\to\infty),$$

即 $u_n\xrightarrow{L^p}u$. 而对每个 $n\in\mathbf{N}$，u_n 看作 $L^p[-n,n]$ 中的元可用 $[-n,n]$ 上的阶梯函数 L^p 逼近. 于是结合(iii)得出，$L^p(\mathbf{R})$ 有如下基本集(与式(1.3.7)对照)：

$$\{\chi_\delta:\delta\subset\mathbf{R} \text{ 是有限区间}\}. \tag{1.3.8}$$

(v) 空间 $L^p(\Omega)$ 的基本集，$1\leqslant p<\infty$，$\Omega\subset\mathbf{R}^n$ 是任意开集. 任给实或复值函数 u，约定

$$\mathrm{supp}u=\overline{\{x:u(x)\neq 0\}}, \tag{1.3.9}$$

称它为函数 u 的**支集**. 令

$$C_c^m(\Omega)=\{u\in C^m(\Omega):\mathrm{supp}u \text{ 是 } \Omega \text{ 的有界子集}\}, \tag{1.3.10}$$

其中 $0\leqslant m\leqslant\infty$. 一个可庆幸的事实是：$C_c^m(\Omega)$ 在 $L^p(\Omega)$ 中稠密，因而是 $L^p(\Omega)$ 的基本集. 当然，我们不会要求你在此处知道上述结论的严格依据. 但你一定应当知道，p 次可积函数(其连续性可能极差)总可以用极好的函数(例如

$C_c^\infty(\Omega)$ 中的函数) L^p 逼近;或者说,**两极相通**,极坏的函数可以用极好的函数 L^p 逼近.读者将有机会体验这一结论的价值.

一般说来,为充分利用基本集的好处,通常应将它取得尽可能地"小",而且其中的元素足够简单,或已被充分研究而容易把握,因而容易在其上建立某些命题;而通过一个极限过程,就可将这些命题推广到全空间上.这种方法在泛函分析的应用中具有基本意义,其显著效果从下面的简单例子能得到初步体现.

例 1.3.9　(i) 设 $u \in L^1(J), J = [a,b]$, 今证

$$\lim_n \int_a^b u(x)\sin nx\,dx = 0. \tag{1.3.11}$$

若 $u = \chi_{(\alpha,\beta)}, (\alpha,\beta) \subset J$, 则显然

$$\int_a^b u(x)\sin nx\,dx = \int_\alpha^\beta \sin nx\,dx \to 0 \quad (n \to \infty).$$

由此容易推出,当 $u(\cdot)$ 是阶梯函数时式(1.3.11)也成立.现在考虑一般的 $u \in L^1(J)$. $\forall \varepsilon > 0$, 取 J 上的阶梯函数 v, 使 $\|u - v\|_1 < \varepsilon$. 取 $N > 0$, 使得

$$\left|\int_a^b v(x)\sin nx\,dx\right| < \varepsilon \quad (\forall n \geqslant N).$$

于是当 $n \geqslant N$ 时有

$$\begin{aligned}&\left|\int_a^b u(x)\sin nx\,dx\right| \\ \leqslant&\int_a^b |u(x) - v(x)|\,dx + \left|\int_a^b v(x)\sin nx\,dx\right| \\ <&\|u - v\|_1 + \varepsilon < 2\varepsilon,\end{aligned}$$

这正表明式(1.3.11)成立.

(ii) 设 $u \in L^1(\mathbf{R})$, $u(x)$ 的 Fourier 变换定义为

$$\hat{u}(\omega) = \int_{-\infty}^{\infty} u(x)e^{-i\omega x}\,dx.$$

今证

$$\hat{u}(\pm\infty) = \lim_{\omega \to \pm\infty} \hat{u}(\omega) = 0. \tag{1.3.12}$$

若 $u = \chi_{(\alpha,\beta)}$, 则

$$\hat{u}(\omega) = \int_\alpha^\beta e^{-i\omega x}\,dx = \frac{e^{-i\omega\beta} - e^{-i\omega\alpha}}{-i\omega} \to 0 \quad (|\omega| \to \infty).$$

由此又推出,当 u 是有有界支集的阶梯函数时式(1.3.12)成立.若 $u \in L^1(\mathbf{R})$ 是任给的, $\forall \varepsilon > 0$, 取有有界支集的阶梯函数 v, 使 $\| u - v \|_1 < \varepsilon$. 取 $A > 0$, 使得

$$|\hat{v}(\omega)| < \varepsilon \quad (|\omega| \geqslant A).$$

则当 $|\omega| \geqslant A$ 时有

$$\begin{aligned} |\hat{u}(\omega)| &\leqslant |\hat{u}(\omega) - \hat{v}(\omega)| + |\hat{v}(\omega)| \\ &< \| u - v \|_1 + \varepsilon < 2\varepsilon, \end{aligned}$$

这表明式(1.3.12)成立.

以上两例都涉及用阶梯函数 L^1 逼近可积函数.一般说来,对于涉及 L^p 空间 $(1 \leqslant p < \infty)$ 的问题,运用适当的逼近方法最见效果.因 L^p 空间中的函数一般"很坏",如果不挑选一些"较好"的函数去逼近,就可能有难以克服的困难.

§1.4 紧性与纲定理

在数学论证中,经常要遇到以下两类存在性问题:

(A) 确立某种对象(如某个问题的解)存在,即至少存在一个;

(B) 确立某种对象普遍地存在,即存在且充分多.

本节所考虑的紧性与纲定理,将分别为解决问题(A)与(B)提供有力工具.

以下设 X, Y 等是给定的赋范空间.

§1.4.1 紧性

许多存在性论证依赖于以下事实:欲证其存在的对象是某一点集的聚点,而这种聚点的存在性常常又取决于点集中的序列恒存在收敛子列.这就表明,其中任何序列存在收敛子列的点集在存在性论证中可起关键作用.这导致如下定义.

定义 1.4.1 设 $A \subset X$. 若 A 中任何序列有收敛子列,则称 A 为**相对紧集**;若 A 中任何序列有收敛子列且其极限属于 A, 则称 A 为**紧集**.

直接由定义看出,紧集就是闭的相对紧集; A 是相对紧集$\Leftrightarrow \bar{A}$ 是紧集;相对紧集的任何子集是相对紧集;紧集的任何闭子集是紧集.若 $A \subset X$ 无界,则必有序列 $\{x_n\} \subset A$, 使 $\| x_n \| \to \infty (n \to \infty)$, 显然这样的 $\{x_n\}$ 不可能有收敛子列,因而 A 不是相对紧集.这就表明,相对紧集必为有界集,紧集必为有界闭集.数学分析中的一条基本定理指出:$\mathbf{R}$ 中任何有界序列必有收敛子列.这等于断言:$\mathbf{R}$ 中的有界集是相对紧集,从而有界闭集是紧集.只需简单的论证,就可将以上结论推广

于 $\mathbf{R}^n$ 以至任何有限维赋范空间.

定理 1.4.2 设 $\dim X < \infty, A \subset X$. 则 A 是相对紧集 $\Leftrightarrow A$ 有界; A 是紧集 $\Leftrightarrow A$ 是有界闭集.

虽然这一结论并不适用于无限维空间(我们将很快指出这一点,参看稍后的定理 1.4.8),但它能给我们一种初步的直观印象:原来紧集就是类似于 $\mathbf{R}^n$ 中有界闭集(如闭区间)的点集. 当然这一理解还远不足以构成对紧集的准确认识,甚至可能产生误导(参看后面的定理 1.4.8). 但一个明显的有益启示是:既然紧性定义密切地联系于"有界序列有收敛子列"这一分析基本定理,那么像区间套定理、有限覆盖定理这样的分析基本定理是否也应与紧集有某种联系? 这正是以下定理要解答的.

定理 1.4.3 对于 $A \subset X$ 给出以下条件:

(i) A 是紧集;

(ii) A 的任何无限子集必有聚点;

(iii) 若 $A \supset B_n \supset B_{n+1}, B_n (n = 1,2,\cdots)$ 是非空闭集,则 $\bigcap B_n \neq \varnothing$;

(iv) A 的任何开覆盖有有限子覆盖;

(v) A 是**全有界集**,这意味着 $\forall r > 0$, A 可用有限个半径为 r 的球覆盖.

则当 A 相对紧时条件(v)满足,即相对紧集是全有界集;当 A 是闭集时(i)$\Leftrightarrow$(ii)$\Leftrightarrow$(iii)$\Leftrightarrow$(iv)$\rightarrow$(v);当 X 完备时(v)$\Leftrightarrow$ A 相对紧;当 X 完备且 A 为闭集时条件(i)~(v)互相等价.

定理的证明较长,不拟给出,有兴趣的读者可参看文献[10] §1.9. 现在只对定理 1.4.3 中的条件作些解释. 若 $\mathscr{U}$ 是 X 的一族子集,其中诸集之并包含 A,则说 $\mathscr{U}$ 覆盖 A,或称 $\mathscr{U}$ 为 A 的一个覆盖;由开集组成的覆盖就称为**开覆盖**. 若 $\mathscr{V} \subset \mathscr{U}$ 且 $\mathscr{V}$ 覆盖 A,则称 $\mathscr{V}$ 为 $\mathscr{U}$ 的子覆盖;由有限个集组成的子覆盖称为**有限子覆盖**. 与数学分析中的熟知结果相对照,可以说条件(iv)相当于断言:对于集 A 有限覆盖定理成立;条件(iii)不妨称之为**闭集套定理**,它显然可看作区间套定理的推广;至于(ii),自然就是**聚点原理**了. 这样,通过定理 1.4.3,就在抽象空间的一般框架内实现了有限覆盖定理、闭集套定理与聚点原理的统一,而且统一于紧集概念. 鉴于此,紧集概念的理论价值,也就不言而喻了.

不过,我们更关心紧集的应用,现在是该给出某些应用例子的时候了.

§1.4.2 紧性的某些应用

从定理 1.4.3 我们能作出如下结论:凡能用区间套定理、有限覆盖定理等分析基本定理证明的结论,亦可能利用紧性在一般赋范空间中建立起来. 首先容易想到的是,紧集上的连续映射应具有类似于闭区间上连续函数的性质. 这就是下面的定理.

定理 1.4.4 设 $D \subset X$ 是紧集，$F \in C(D,Y)$，$f \in C(D,\mathbf{R})$. 则有以下结论：

(i) FD 是 Y 中的紧集，因而有界；F 在 D 上一致连续，即 $\forall \varepsilon > 0, \exists \delta > 0$，当 $x, y \in D$，$\|x - y\| < \delta$ 时 $\|Fx - Fy\| < \varepsilon$；

(ii) f 在 D 上取得最大值与最小值.

证 (i) 任给序列 $\{x_n\} \subset D$，由 D 紧必有收敛子列 $\{x_{n_k}\}: x_{n_k} \to x \in D$，因而 $Fx_{n_k} \to Fx \in FD$. 这表明 FD 是紧集. 若 F 在 D 上非一致连续，则对某个 $\varepsilon > 0$ 有 $x_n, y_n \in D(n = 1,2,\cdots)$，使得 $x_n - y_n \to 0$，而 $\|Fx_n - Fy_n\| \geqslant \varepsilon(n = 1, 2,\cdots)$（写出"非一致连续"的这一刻画对于证明是关键的）. 如同上面一样，$\{x_n\}$ 有收敛子列 $\{x_{n_k}\}: x_{n_k} \to x \in D(k \to \infty)$；同样有 $y_{n_k} \to x$，但这推出

$$\varepsilon \leqslant \lim_k \|Fx_{n_k} - Fy_{n_k}\| = \|Fx - Fx\| = 0,$$

得出矛盾. 因此 F 必定一致连续.

(ii) 不妨只证最大值存在. 令 $\beta = \sup\limits_{x \in D} f(x)$. 取 $\{x_n\} \subset D$，使得 $f(x_n) \to \beta(n \to \infty)$. $\{x_n\}$ 有收敛子列 $\{x_{n_k}\}: x_{n_k} \to x \in D(k \to \infty)$，则

$$f(x) = \lim_k f(x_{n_k}) = \beta,$$

显然 β 即为 f 在 D 上的最大值. □

对于上述定理证明之简洁，你大概会有深刻印象. 这正是紧集概念优点之所在. 实际上，我们甚至可将定理 1.4.4 的证明表述得更简单些. 例如在其(ii)的证明中，只需改变记号，就可用 $\{x_n\}$ 取代 $\{x_{n_k}\}$. 因此可将证明改述如下：取 $\{x_n\} \subset D$，使 $f(x_n) \to \beta(n \to \infty)$. 不妨设 $x_n \to x \in D$，则 $f(x_n) \to f(x) = \beta$，因此 β 即为 f 之最大值.

今后在类似的论证中，经常用"不妨"二字来避免直接提到子列. 这种简化笔法，是值得你细加体会的.

若 $\bar{x} \in D$ 使得 $f(\bar{x}) = \min\limits_{x \in D} f(x)$，则通常说 $\bar{x}$ 是**最小化问题**

$$\min f(x), \quad x \in D \tag{1.4.1}$$

的**最优点**或**最优解**. 于是定理 1.4.1 之(ii)无非是说，若 D 为紧集且 $f \in C(D,\mathbf{R})$，则问题(1.4.1)式的最优解存在. 类似于式(1.4.1)的问题大量地出现于各个领域，有关的存在性论证往往要在这样或那样的形式下运用紧性. 下面是另一个例子.

推论 1.4.5(最佳逼近) 设 A 是 X 的有限维子空间，$x \in X$. 则存在 $a \in A$，使得 $\|x - a\| = d(x,A)$. 换言之，a 是 A 中离 x 最近的点，因而称为 x 在 A 中的**最佳逼近**.

事实上，只要取 $r > 0$ 充分大，x 的最佳逼近 a 就是最小化问题

$$\min \| x - a \|, \quad a \in A \cap \overline{B}_r(0) \tag{1.4.2}$$

的最优解(何故?).由定理1.4.2,有界闭集 $A \cap \overline{B}_r(0)$ 是紧集,故所要结论由定理1.4.4(ii)推出.如果不用定理1.4.4,用一直接证明也很简单:取 $\{a_n\} \subset A$，使得 $\| x - a_n \| \to d(x,A)(n \to \infty)$. 由

$$\| a_n \| \leqslant \| x \| + \| a_n - x \|$$

知 $\{a_n\}$ 有界.因 $\dim A < \infty$，不妨设 $a_n \to a \in A$，显然 $\| x - a \| = d(x,A)$.

若令 A 为次数 $\leqslant n$ 的多项式之全体,取 $X = C(J)$ (或 $X = L^p(J), 1 \leqslant p < \infty$),则由定理1.4.5立即得到:任给 $u \in X$，必存在次数 $\leqslant n$ 的多项式 v，它是对 u 的最佳一致(或 L^p)逼近.这种存在性结论,是深刻而有价值的,遗憾的只是定理1.4.5并未提供求最佳逼近的方法.对于 L^2 逼近问题,将在下节中通过另外的途径予以考虑.

§1.4.3 紧集的判定

鉴于紧集在理论分析中具有难以比拟的优势,寻求各种具体的紧集判别法就成为一重要课题,这一课题仅对于某些特殊的空间得到了令人满意的解决.特别,对于§1.2中讨论的那些常用的函数空间中的紧集,已有方便的判别法.下面介绍几个这样的判别法而略去其证明.必要先说明的两点是:

(i) 因紧集就是闭的相对紧集,故仅需给出相对紧性的判别法;

(ii) 因相对紧集必为有界集,故下面所给出的相对紧条件中,"有界性"是当然的组成部分.

定理 1.4.6(Arzela-Ascoli 定理) $A \subset C(J)$ 相对紧的充要条件是:

(i) A **一致有界**,即 A 依 sup 范数有界;

(ii) A **等度连续**,即 $\forall \varepsilon > 0, \exists \delta > 0, \forall x,y \in J, \forall u \in A$，当 $|x - y| < \delta$ 时恒有 $|u(x) - u(y)| < \varepsilon$.

将闭区间 J 换成任何有界闭区域 $\Omega \subset \mathbf{R}^n$，以上定理仍然保持成立.

一个应用定理1.4.6的简单例子如下:若 $A \subset C^1(J)$ 依范数 $\| \cdot \|_1$ (定义依式(1.2.6))有界,则 A 作为空间 $C(J)$ 的子集是相对紧的.事实上,设 $\beta = \sup\limits_{u \in A} \| u \|_1$，则 $\sup\limits_{u \in A} \| u \|_0 \leqslant \beta$,

$$|u(x) - u(y)| \leqslant \beta |x - y| \quad (x,y \in J, u \in A).$$

可见 A 一致有界且等度连续,因而由定理1.4.6是相对紧的.

定理 1.4.7 设 $1 \leqslant p < \infty, A \subset l^p$，则 A 相对紧的充要条件是:

(i) A 有界,即 $\sup_{x\in A}\|x\|_p<\infty$;

(ii) 关于 $x=(x_i)\in A$ 一致地有 $\sum_{i>n}|x_i|^p\to 0(n\to\infty)$, 即 $\forall\varepsilon>0$, $\exists N>0,\forall n\geqslant N,\forall x\in A$, 有 $\sum_{i>n}|x_i|^p<\varepsilon$.

定理 1.4.7 的一个直接推论是: $A=\{x=(x_i):|x_i|\leqslant 1/i(\forall i\in\mathbf{N})\}$ 是空间 l^2 中的紧集,它就是所谓 **Hilbert 方体**.这是一个结构具体而又较简单的无限维紧集的例子.在对无限维空间中的紧集缺少直接经验的情况下,这样的例子是特别有价值的.

对照定理 1.4.6,定理 1.4.7 与定理 1.4.2 可看出,对于空间 $C(J)$ 与 l^p 中的相对紧性,都在有界性条件的基础上增加了一个条件(ii).只要空间是无限维的,这一类的附加条件就不可缺少.为证实这一点,请看令人惊异的如下定理.

定理 1.4.8 若 $\dim X=\infty$ [1],则 X 中的闭单位球不是紧集.

为证明定理 1.4.8,需要以下引理,它在别处也大有用处.

引理 1.4.9(Riesz 引理) 设 A 是 X 的闭子空间, $A\neq X$, $0<r<1$. 则存在 $x\in X$, 使得 $d(x,A)>r$ 且 $\|x\|=1$.

证 取 $y\in X\setminus A$, 必定 $\rho\triangleq d(y,A)>0$ (用命题 1.3.2(i)).取定 $a\in A$, 使得 $\|y-a\|<\rho/r$ (这种 a 何以存在?).令 $x=(y-a)/\|y-a\|$, 则 $\|x\|=1$,

$$d(x,A)=d\left(\frac{y-a}{\|y-a\|},A\right)=\frac{d(y-a,A)}{\|y-a\|}$$

$$=\frac{d(y,A)}{\|y-a\|}>\rho\cdot\frac{r}{\rho}=r. \qquad\square$$

若 A 是 Euclid 空间 $\mathbf{R}^n$ 的真子空间,则必有单位向量 x, 使得 x 垂直于 A (即 x 垂直于 A 中每个向量),因而 $d(x,A)=1$. 引理 1.4.9 中的 x 似乎接近于有上述性质,因此有人说引理 1.4.9 中的 x "几乎垂直于 A",尽管 X 中一般并无垂直概念.

(定理 1.4.8 的证明) 因 $\dim X=\infty$, 故存在线性无关的无限序列[2] $\{a_n\}\subset X$. 令

$$X_n=\operatorname{span}\{a_1,a_2,\cdots,a_n\},$$

---

[1] 记号 $\dim X=\infty$ 仅表示 X 不是有限维空间,并不涉及是否有一个无限的维数 $\dim X$, 尽管这种维数是可以定义的.

[2] 向量空间中一无限集线性无关,意指它的任何非空有限子集线性无关.

则 X_n 是 X 的闭子空间，且 $X_n \subsetneqq X_{n+1}(n=1,2,\cdots)$. 分别以 X_{n-1}, X_n 代 A, X 应用引理 1.4.9，知存在 $x_n \in X_n$，使得 $d(x_n, X_{n-1}) > 1/2$ 且 $\|x_n\| = 1(n=2,3,\cdots)$. 因显然有

$$\|x_m - x_n\| \geqslant 1/2 \quad (m \neq n, m, n \geqslant 2),$$

故 $\{x_n\}$ 必无收敛子列. 这表明闭球 $\overline{B}_1(0)$ 不是紧集. □

注意以上论证实际上证明了无限维赋范空间中的单位球面 $S \triangleq \{x : \|x\| = 1\}$ 不是紧集，因为其上存在无聚点的无限子集. 如果囿于平常的直观，你会惊讶于单位球面上何以容得下无聚点的无限集. 试想若将无限个点置于地球表面，它们不总会在某个局部高度聚集吗？问题在于，你的观察局限在 $\mathbf{R}^3$ 中了. 若以 $\{e_i\}$ 记 $\mathbf{R}^n$ 的标准基，则 $2n$ 个点 $\pm e_i(1 \leqslant i \leqslant n)$ 构成一个两两相距 $\sqrt{2}$ 的点集. 随着 n 的增大，这个点集的点数将无限增长，但其稀疏的程度却保持不变. 这就可以想像，当 $\dim X$ 增至无穷时，单位球面将变得如此广袤，以至即使无限多个点分布于其上，它们也可能呈离散状态而无聚点，就如同定理 1.4.8 证明中的 $\{x_n\}$ 一样.

因任何(半径>0)闭球可通过平移与相似变换①而互相变换，而这些变换并不改变紧性(何故?)，故定理 1.4.8 推出无限维赋范空间中任何闭球是非紧的. 这显然又推出，无限维赋范空间中任何含内点的集是非紧的，因而紧集必无内点②. 这就多少揭示了无限维赋范空间中紧集的独特性质. 以上分析也表明，在某种意义上，无限维赋范空间中紧集甚为稀少，这通常是处理无限维问题的困难之所在. 例如，设 S 是 X 中的单位球面，$f \in C(S, \mathbf{R})$，则当 $\dim X < \infty$ 时可断定 f 在 S 上取得最大值与最小值，而当 $\dim X = \infty$ 时已不能作此断言！下面就是一个简单的反例. 设 $X = C[0,1]$，

$$f(u) = \int_0^1 |u(x)| \, \mathrm{d}x \quad (u \in X).$$

显然 $f \in C(X, \mathbf{R})$，且在 $S \triangleq \{u \in X : \|u\|_0 = 1\}$ 上 $f(u) > 0$. 取 $u_n(x) = x^n (0 \leqslant x \leqslant 1)$，则 $u_n \in S$，$f(u_n) = 1/(n+1) \to 0 (n \to \infty)$，而这就表明 $\inf\limits_{u \in S} f(u) = 0$，$f$ 在 S 上取不到最小值.

§1.4.4　纲定理

本节开头提出的问题(B)，似乎更具有挑战性，初看起来难以想像有什么好的

---

① 对固定的 x_0 与 $0 \neq \alpha \in \mathbf{K}$，$X$ 到自身的映射 $x \to x + x_0$ 与 $x \to \alpha x$ 分别称为平移与相似变换，二者都连续且有连续的逆映射.

② 特别，Hilbert 方体 $A = \{x = (x_i) \in l^2 : |x_i| \leqslant 1/i (\forall i \in \mathbf{N})\}$ 必无内点. 你会认为 $\{(x_i) \in l^2 : |x_i| < 1/i (\forall i \in \mathbf{N})\}$ 是开集，而 $x = 0$ 是 A 的内点吗？

解决办法,而且甚至不知道该如何严格地描述某个对象**充分多**.我们首先得界定,若 $A\subset X$,当 A 具有什么特征时可以说,任何元 $x\in X$ **一般地**在 A 中,或者相反,**一般地**不在 A 中.实际上,可行的方案不只一种,下面给出的纲概念或许提供了一种最有效的办法.

定义 1.4.10 设 $A\subset X$. 若 $(\bar{A})^{\circ}=\varnothing$, 则称 A 为**疏集**.可数个疏集之并称为**第一纲集**;非第一纲集称为**第二纲集**;第一纲集的补集称为**剩余集**.

直接由定义看出,无内点的闭集(例如 $\mathbf{R}^2$ 中的直线)是疏集.特别,单点集总是疏集,因而可数集是第一纲集;例如 $\mathbf{R}$ 中的有理点集就是第一纲集.当然,疏集也是第一纲集.由此推想,第一纲集似乎所含点"较少",是一种**瘦集**,可用来描述**稀有的**对象;而第二纲集则是**非瘦集**.进一步的描述已超出通常直观可及的视野.不过,从逻辑上说,有一点是明显的:要使用第一纲集描述稀有性具有合理性,X 本身必须是第二纲集(这等价于剩余集是第二纲集),因而如下的纲定理具有基本的重要性.

定理 1.4.11(Baire 纲定理) 设 X 完备,$A\subset X$ 是第一纲集,则 A^c 是第二纲集且为稠集.

证 首先证 A^c 是稠集,即 $\overline{A^c}=X$, 为此只需证 $A^{\circ}=\varnothing$ (用式(1.3.3b)).用反证法:若 $A^{\circ}\neq\varnothing$, 则 A° 是非空开集.设 $A=\bigcup_1^{\infty}A_n$,$A_n(n=1,2,\cdots)$ 是疏集.由 $(\bar{A}_1)^{\circ}=\varnothing$ 推出 $(\bar{A}_1)^c$ 是开的稠集(再用式(1.3.3b)),故

$$A^{\circ}\setminus\bar{A}_1=A^{\circ}\cap(\bar{A}_1)^c$$

是非空开集,因而其中必含一个闭球 $\bar{B}_{r_1}(x_1)$, 可设 $0<r_1<1$. 分别以 $B_{r_1}(x_1)$ 与 A_2 替换 A° 与 A_1, 重复以上论证得

$$\bar{B}_{r_2}(x_2)\subset B_{r_1}(x_1)\setminus\bar{A}_2,\quad 0<r_2<1/2.$$

一般地,可递次得出

$$\bar{B}_{r_n}(x_n)\subset B_{r_{n-1}}(x_{n-1})\setminus\bar{A}_n,\quad 0<r_n<1/n,$$

$(n=2,3,\cdots)$. 当 $m>n$ 时显然 $\bar{B}_{r_m}(x_m)\subset B_{r_n}(x_n)$, 这推出

$$\|x_m-x_n\|<r_n<1/n,\quad m>n=1,2,\cdots$$

故 $\{x_n\}$ 是 Cauchy 列.由完备性,有 $x_n\to x\in X(n\to\infty)$. 但

$$x\in\bar{B}_{r_n}(x_n)\subset B_{r_{n-1}}(x_{n-1})\setminus\bar{A}_n\subset A^{\circ}\setminus\bar{A}_1\subset A\quad(n\geqslant 2),$$

这推出 $x\in A$ 而 $x\overline{\in}\bigcup_1^{\infty}A_n$, 得出矛盾.因此 $A^{\circ}=\varnothing$.

其次证 A^c 为第二纲集.若 A^c 是第一纲集,则 $X=A\cup A^c$ 是第一纲集,但由

已证结论这将推出 $X^\circ = \varnothing$，得出矛盾. □

鉴于 Banach 空间中的剩余集是第二纲集且为稠集，而且，从其中除去任意可数多个第一纲集之后仍然是剩余集，因而仍保持为第二纲集与稠集. 这充分显示出剩余集的某种**不可穷竭性**，因而是描述一般性的适当工具. 为了达到更形象的效果，不妨采用如下说法：若 A 是 Banach 空间 X 中的剩余集，就说几乎每个 $x \in X$ 都属于 A. 例如，无理点集是 $\mathbf{R}$ 中的剩余集，因而几乎每个实数都是无理数！

应用纲定理的途径有两条. 一是利用基于 Baire 定理的某些标准结果，例如下章将介绍的逆算子定理与一致有界原理；二是直接应用 Baire 定理. 下面举一个直接应用的例子.

例 1.4.12 $J = [a, b]$ 上几乎所有连续函数处处不可微.

证 设 $X = C(J)$. 令

$$A = \{u \in X : u \text{ 在}[a, b) \text{ 上至少一点处可微}\};$$

$$B = \{u \in X : u \text{ 在}(a, b] \text{ 上至少一点处可微}\}.$$

显然 $(A \cup B)^c$ 正是 J 上处处不可微的连续函数之全体，因而只要证 A 与 B 均为 X 中的第一纲集. 不妨只证 A 是第一纲集. 令

$$A_n = \left\{u \in X : \exists x \in \left[a, b - \frac{1}{n}\right]\right.$$

使

$$\left.\sup_{h>0} \frac{|u(x+h) - u(x)|}{h} \leqslant n\right\},$$

则不难看出 $A \subset \bigcup_1^\infty A_n$，因此只需证 $A_n (n = 1, 2, \cdots)$ 是 X 中的疏集. 取定 $n \in \mathbf{N}$. 先证 A_n 是闭集. 若 $\{u_k\} \subset A_n$，$u_k \rightrightarrows u (k \to \infty)$，则有 $x_k \in [a, b - n^{-1}]$，使得

$$|u_k(x_k + h) - u_k(x_k)| \leqslant nh \quad (x_k < x_k + h \leqslant b). \tag{1.4.3}$$

不妨设 $x_k \to x \in [a, b - n^{-1}] (k \to \infty)$. 令 $k \to \infty$ 由式(1.4.3)得

$$|u(x+h) - u(x)| \leqslant nh \quad (x < x + h \leqslant b),$$

这表明 $u \in A_n$，即 A_n 是闭集. 余下只要证 $A^\circ_n = \varnothing$，即 $\overline{A_n^c} = X$（用式(1.3.3b)）. 注意

$$A_n^c = \left\{u \in X : \forall x \in \left[a, b - \frac{1}{n}\right], \text{有} \sup_{h>0} \frac{|u(x+h) - u(x)|}{h} > n\right\},$$

由此不难看出，一个如图 1-3 之左图所示的振动足够快的“锯齿形函数”必属于

A_n^c. 显然每个 $u \in X$ 都可以用如上的函数一致逼近,因此 $\overline{A_n^c} = X$, 如所要证 □

例 1.4.12 的结论初看起来颇令人惊异. 如果考虑到,构造一个处处不可微的连续函数并不是一件平凡的事(这样的例子最早由 Weierstrass 于 1872 年给出,当时在数学界引起的震动是可想而知的),而可微函数则似乎俯拾皆是,那么,例 1.4.12的结论就更加惊人了. 概言之,可微性与连续性的差别比初看起来要大得多. 直观上,这意味着连续曲线很少是光滑的. 无疑,任何具有足够数学修养的现代数学工作者,即使在直观上也不会再排斥上述结论,但却不大可能仅凭初等的考虑达到严格的理解. 这就使泛函分析所提供的抽象空间方法有了充分的用武之地.

§1.5 Hilbert 空间

在前几节中,我们实际上将 Euclid 空间中处理问题的一些思路移植到了 Banach空间. 通过自然的类比并借用平常的几何语言,Banach 空间理论呈现出某种可加直观想像的面貌. 这样作的效果如此引人入胜,人们不禁想再往前走一步:将更多的几何概念,如角度、垂直性等,都引进到抽象空间中来. 这些概念中,最重要的是垂直性概念. 而从解析几何我们知道,垂直性与内积有关. 因此,关键性的一步是推广内积概念,而这导向一般的内积空间与 Hilbert 空间理论.

§1.5.1 内积空间

从线性代数中熟知, $\mathbf{K}^n$ 中由如下公式定义了内积:

$$\langle x, y\rangle = \sum_i x_i \bar{y}_i, \quad (x = (x_i), y = (y_i) \in \mathbf{K}^n) \tag{1.5.1}$$

且

$$|x| = \sqrt{\langle x, x\rangle} = \Big(\sum_i |x_i|^2\Big)^{1/2} \tag{1.5.2}$$

正是 Euclid 范数(参照式(1.1.1)). §1.1.1 中引进范数的经验启示我们,内积概念的一般化不是去推广公式(1.5.1),而是依赖于概括内积的某些性质. 这一考虑导致如下的公理化定义.

定义 1.5.1 设 H 是 $\mathbf{K}$ 上的向量空间. 若对任一对元 $x, y \in H$, 指定了一个数 $\langle x, y\rangle \in \mathbf{K}$, 称为 x 与 y 的**内积**①,它满足以下**内积公理**:

(I_1) $\langle\cdot, y\rangle$**的线性性**: $\langle \alpha x + \beta z, y\rangle = \alpha\langle x, y\rangle + \beta\langle z, y\rangle$;

---

① 内积一词,今后将兼指特定向量 x 与 y 的内积 $\langle x, y\rangle$ 与二元函数 $(x, y) \to \langle x, y\rangle$; 在后一种意义上使用时,通常写作 $\langle\cdot, \cdot\rangle$. 参照第 2 页注①.

(I_2) **共轭对称性**：$\langle x,y\rangle=\overline{\langle y,x\rangle}$；

(I_3) **正定性**：$\langle x,x\rangle\geqslant 0$；$\langle x,x\rangle=0\Leftrightarrow x=0$，

(以上 $x,y,z\in H,\alpha,\beta\in\mathbf{K}$)，则称 H 为 $\mathbf{K}$ 上的**内积空间**；当要明确指出内积时将 H 写作 $(H,\langle\cdot,\cdot\rangle)$. 若 $\mathbf{K}=\mathbf{R}$(或 $\mathbf{C}$)，则称 $\mathbf{K}$ 上的内积空间为实(或复)内积空间.

显然，$\mathbf{K}^n$ 依式(1.5.1)所定义的内积就是一个内积空间，称内积(1.5.1)式为 $\mathbf{K}^n$ 上的标准内积，也写作 $x\cdot y$ 或 $x^T\bar{y}$. $\mathbf{K}^n$ 常用作解释内积空间概念的模型.

以下设 H 是 $\mathbf{K}$ 上给定的内积空间.

我们首先看公理(I_1)～(I_3)有哪些推论. 结合(I_1)与(I_2)知 $\langle x,\cdot\rangle$ 是**共轭线性**的：

$$\langle x,\alpha y+\beta z\rangle=\bar{\alpha}\langle x,y\rangle+\bar{\beta}\langle x,z\rangle\quad(x,y,z\in H,\alpha,\beta\in\mathbf{K}).$$

这与(I_1)合起来表明，内积 $\langle x,y\rangle$ 是"一个半线性"的：对 x 是线性的，对 y 是共轭线性的. 若 $\mathbf{K}=\mathbf{R}$，则内积是**双线性**的，而公理 I_2 等价于**对称性**：$\langle x,y\rangle=\langle y,x\rangle$. 一般地，若 $\sum\alpha_i x_i$ 与 $\sum\beta_j y_j$ 是 H 中元的两个有限线性组合，则

$$\langle\sum_i\alpha_i x_i,\sum_j\beta_j y_j\rangle=\sum_{i,j}\alpha_i\bar{\beta}_j\langle x_i,y_j\rangle;\tag{1.5.3}$$

$$\langle\sum_i\alpha_i x_i,\sum_i\alpha_i x_i\rangle=\sum_{i,k}\alpha_i\bar{\alpha}_k\langle x_i,x_k\rangle.\tag{1.5.4}$$

可见，内积运算的规则大体上相当于通常的乘法规则，只是记住在适当位置上加上共轭符号.

约定 $\|x\|=\sqrt{\langle x,x\rangle}\,(x\in H)$. 从式(1.5.2)看来，$\|x\|$ 应是一个范数，下面验明它确实满足范数公理(N_1)～(N_3)(依定义 1.1.1). 公理(N_1)，(N_3)的验证是直接的. 为验证(N_2)，先建立一个不等式，它本身有重要的独立价值.

引理 1.5.2(Schwarz 不等式)　对任给 $x,y\in H$，成立

$$|\langle x,y\rangle|\leqslant\|x\|\,\|y\|.\tag{1.5.5}$$

证　取定 $x,y\in H$. 对任给 $\alpha\in\mathbf{K}$，有

$$\begin{aligned}0&\leqslant\langle x-\alpha y,x-\alpha y\rangle &&(\text{用}(I_3))\\&=\|x\|^2-\alpha\langle y,x\rangle-\bar{\alpha}\langle x,y\rangle+\alpha\bar{\alpha}\|y\|^2.&&(\text{用式}(1.5.4))\end{aligned}$$

可设 $y\neq 0$(否则式(1.5.5)已成立). 取 $\alpha=\langle x,y\rangle/\|y\|^2$ 代入以上不等式，整理后恰得不等式(1.5.5). □

现在利用 Schwarz 不等式来验证三角不等式. 任给 $x,y\in H$，有

$$\begin{aligned}\|x+y\|^2 &= \langle x+y, x+y\rangle \\ &= \|x\|^2 + \langle x,y\rangle + \langle y,x\rangle + \|y\|^2 \quad (\text{用}(1.5.4)\text{式}) \\ &\leqslant \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2 \quad (\text{用}(1.5.5)\text{式}) \\ &= (\|x\| + \|y\|)^2,\end{aligned}$$

这得 $\|x+y\| \leqslant \|x\| + \|y\|$. 这就验证了 $\|x\| = \sqrt{\langle x,x\rangle}$ 确为 H 上的范数,称它为由内积定义的范数.在未加说明时,在内积空间中总使用由内积定义的范数;依此,内积空间自动地被看作赋范空间,当其完备时称为 **Hilbert 空间**;**R**(或**C**)上的 Hilbert 空间称为实(或复)Hilbert 空间①.

内积空间作为特殊的赋范空间,当然具有一般赋范空间的所有性质.我们所关注的是内积空间所特有的那些性质.首先指出一个简单而常用的结论:内积依范数收敛是连续的,即若在 H 中 $x_n \to x, y_n \to y$, 则 $\langle x_m, x_n\rangle \to \langle x,y\rangle (m,n\to\infty)$, 即

$$\lim_{m,n}\langle x_m, y_n\rangle = \langle \lim_m x_m, \lim_n y_n\rangle, \tag{1.5.6}$$

或者说,内积的极限等于极限的内积.这由如下明显的推理得出:

$$\begin{aligned}|\langle x_m, y_n\rangle - \langle x,y\rangle| &\leqslant |\langle x_m,y_n\rangle - \langle x_m, y\rangle| + |\langle x_m,y\rangle - \langle x,y\rangle| \\ &\leqslant \|x_m\|\,\|y_n - y\| + \|x_m - x\|\,\|y\|. \quad (\text{用}(1.5.5)\text{式})\end{aligned}$$

Hilbert 空间的典型例子是空间 $L^2(\Omega)$ ②,此处 (Ω,μ) 是任一测度空间.在 $L^2(\Omega)$ 中定义内积:

$$\langle u,v\rangle = \int_\Omega u(x)\overline{v(x)}\mathrm{d}\mu \quad (u,v\in L^2(\Omega)). \tag{1.5.7}$$

直接看出内积(1.5.7)式满足内积公理$(\mathrm{I}_1)\sim(\mathrm{I}_3)$,且 $\sqrt{\langle u,u\rangle}$ 正是 u 的 L^2 范数.因此, $L^2(\Omega)$ 是一个 Hilbert 空间. $L^2(\Omega)$ 中的 Schwarz 不等式就是 $p=q=2$ 时的 Hölder 不等式(亦称 Cauchy 不等式):

$$\left|\int_\Omega u(x)v(x)\mathrm{d}\mu\right|^2 \leqslant \int_\Omega |u(x)|^2\mathrm{d}\mu\int_\Omega |v(x)|^2\mathrm{d}\mu \quad (u,v\in L^2(\Omega)).$$

---

① 就本节的内容而言,实 Hilbert 空间与复 Hilbert 空间并无重大区别.不过,对于线性算子问题来说,只有在复 Hilbert 空间上,才能建立一个充分完全的理论(参看后面的 §3.4–§3.5).

② 在某种完全严格的意义上,适当选择 Ω, 可使 $L^2(\Omega)$ 穷尽所有 Hilbert 空间.因此可以说,本质上, **$L^2(\Omega)$ 是仅有的 Hilbert 空间**.空间 L^2 之重要,由此可见.

类似地，l^2 亦为 Hilbert 空间，其中的内积定义为

$$\langle x,y\rangle=\sum_i x_i\bar{y}_i\quad (x=(x_i),y=(y_i)\in l^2). \tag{1.5.8}$$

形式上，式(1.5.1)与(1.5.8)几乎一致. 在 l^2 中，Schwarz 不等式成为

$$\left|\sum_i x_i\bar{y}_i\right|^2\leqslant\sum_i|x_i|^2\sum_i|y_i|^2$$

$((x_i),(y_i)\in l^2)$，这实际上就是 Cauchy 不等式(1.2.11).

你注意到，此处未提到 §1.2 中所述的其他大多数空间，如 $L^p,l^p(p\neq 2)$，$C^m(\Omega)$ 等. 实际上，所有这些空间都不是内积空间. 有一个很简单的规则用来检验：给定赋范空间是否为内积空间(即其范数是否可由一内积定义). 这就是下面的定理.

定理 1.5.3 $\mathbf{K}$ 上的赋范空间 X 是内积空间的充要条件是，其中的范数满足如下**中线公式**：

$$\|x+y\|^2+\|x-y\|^2=2(\|x\|^2+\|y\|^2). \tag{1.5.9}$$

证 若 X 是内积空间，以 $\|x\|^2=\langle x,x\rangle$ 代入(1.5.9)式很容易验证其为等式. 反之，若式(1.5.9)成立，当 $\mathbf{K}=\mathbf{R}$ 与 $\mathbf{K}=\mathbf{C}$ 时分别令

$$4\langle x,y\rangle=\|x+y\|^2-\|x-y\|^2 \tag{1.5.10}$$

与

$$4\langle x,y\rangle=\|x+y\|^2-\|x-y\|^2+\mathrm{i}\|x+\mathrm{i}y\|^2-\mathrm{i}\|x-\mathrm{i}y\|^2, \tag{1.5.10$'$}$$

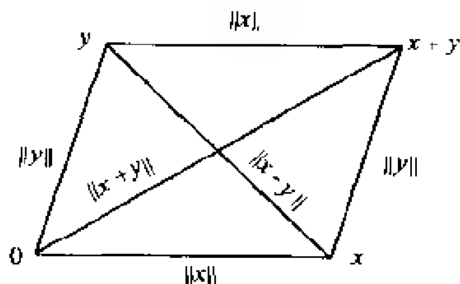


图 1-6

则可验证 $\langle x,y\rangle$ 是内积且 $\langle x,x\rangle=\|x\|^2$，因而 X 是内积空间. 验证细节是技术性的，可参看文献[4]第 515 页. □

公式(1.5.10)与(1.5.10)′称为**极化恒等式**，用处颇大，凡需要从范数的一定结论推出内积的相应结论时，就应想到**极化恒等式**.

中线公式(1.5.9)正好表达了熟知的几何定理：平行四边形各边平方和等于两对角线平方和(见图 1-6). 注意，$\|x+y\|$ 正是以 $0,x,y$ 为顶点的三角形之一中线的两倍.

§1.5.2 正交系

现在回到一般的内积空间 H. 内积空间区别于一般赋范空间的主要特点是其

中可定义正交性. 利用正交性概念,就可将直角坐标、垂直投影等一系列 Euclid 几何概念推广于内积空间,从而建立内积空间几何学. 这就使 Hilbert 空间与 Euclid 空间的类比达到近乎完美的程度. 利用内积,正交性的定义是很简单的.

定义 1.5.4 (i) 设 $x,y\in H$. 若 $\langle x,y\rangle=0$, 则说 x 与 y **正交**或**直交**,记作 $x\perp y$.

(ii) 设 $\{x_i:i\in I\}\subset H$. 若当 $i\neq j$ 时 $x_i\perp x_j$, 则称 $\{x_i\}$ 为**正交系** (或正交集、正交组). 若 $\{x_i\}$ 是正交系且 $\|x_i\|=1$ (这等价于 $\langle x_i,x_j\rangle=\delta_{ij}$, δ_{ij} 是熟知的 **Kronecker 记号**),则称 $\{x_i\}$ 为**标准正交系**.

(iii) 设 $A,B\subset H$. 约定 $A\perp B\Leftrightarrow\forall a\in A,\forall b\in B$, 有 $a\perp b$; $x\perp A\Leftrightarrow\forall a\in A$, 有 $x\perp a$; $A^{\perp}=\{x\in H:x\perp A\}$, 称 $A^{\perp}$ 为 A 的**正交补**. 当 $A\perp B$ 时说 A 与 B 互相正交.

若 $\{x_i:1\leqslant i\leqslant n\}$ 是一有限正交系,则直接用(1.5.4)得出

$$\left\|\sum_i x_i\right\|=\sum_i\|x_i\|^2. \tag{1.5.11}$$

特别,取 $n=2$ 得

$$\|x_1\pm x_2\|^2=\|x_1\|^2+\|x_2\|^2.$$

可看出这正是熟知的**勾股定理**(见图 1-7). 因此,可以说式(1.5.11)表达了一般的勾股定理. 以 $\alpha_i x_i$ 代 $x_i(\alpha_i\in\mathbf{K},1\leqslant i\leqslant n)$, 从式(1.5.11)得

$$\left\|\sum_i\alpha_i x_i\right\|^2=\sum_i|\alpha_i|^2\|x_i\|^2. \tag{1.5.12}$$

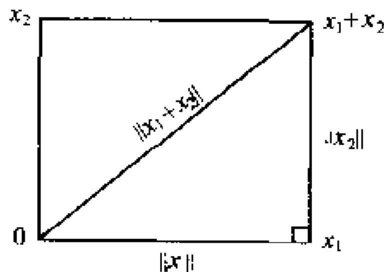


图 1-7

若 $\{x_i\}$ 中无零元,则由式(1.5.12)推出: $\sum\alpha_i x_i=0\Leftrightarrow\alpha_i=0(1\leqslant i\leqslant n)$. 由此可见, $\{x_i\}$ 线性无关,即不含零元的正交系必线性无关.

设 $\{e_i:i\in\mathbf{N}\}$ 是 H 中的标准正交系. 若 $x\in H$ 可表为

$$x=\sum_i\alpha_i e_i, \tag{1.5.13}$$

则

$$\langle x,e_i\rangle=\left\langle\sum_j\alpha_j e_j,e_i\right\rangle=\alpha_i,\quad(\text{用(1.5.6)式})$$

可见表达式(1.5.13)的系数惟一确定. 若每个 $x\in H$ 都可以表如式(1.5.13),则

称 $\{e_i\}$ 为 H 的**标准正交基**. 显然,标准正交基正是与 Euclid 空间中的直角坐标系相当的东西,而式(1.5.13)中的 $\alpha_i = \langle x, e_i\rangle$ 则相当于通常的直角坐标,称为 x 关于基 $\{e_i\}$ 的**正交坐标**. 基本的问题是:标准正交系应具备什么条件才能成为标准正交基? Hilbert 空间是否一定有标准正交基? 这些问题已得到完全的解决,其结论如下.

定理 1.5.5　设 $\{e_i : i \in \mathbf{N}\}$①是 Hilbert 空间 H 中的标准正交系,则以下条件互相等价:

(i) $\{e_i\}$ 是 H 的标准正交基;

(ii) $\{e_i\}$ 是 H 的基本集(依定义 1.3.7(iii));

(iii) $\{e_i\}$ 是极大正交系,即若 $x \perp e_i (\forall i \in \mathbf{N})$, 则 $x = 0$;

(iv) 对任给 $x \in H$, 成立如下 **Parseval 等式**:

$$\| x \|^2 = \sum_i |\langle x, e_i\rangle|^2; \tag{1.5.14}$$

(v) 对任给 $x, y \in H$, 成立如下内积公式:

$$\langle x, y\rangle = \sum_i \langle x, e_i\rangle \overline{\langle y, e_i\rangle}. \tag{1.5.15}$$

证　显然(i)⇒(ii). 为书写简便,下面记 $\hat{x}_i = \langle x, e_i\rangle (x \in H, i \in \mathbf{N})$.

(ii)⇒(iii). 设条件(ii)满足, $x \perp e_i (\forall i \in \mathbf{N})$, 今要证 $x = 0$, 为此只要证 $\| x \|^2 = 0$. 取序列 $\{x_n\} \subset \operatorname{span}\{e_i\}$, 使得 $x_n \to x$. 易见 $x \perp x_n (\forall n \in \mathbf{N})$, 于是

$$\| x \|^2 = \lim_n \langle x, x_n\rangle = 0. \quad (\text{用式}(1.5.6))$$

(iii)⇒(i). 设条件(iii)满足. 取定 $x \in H$, 令 $s_n = \sum_1^n \hat{x}_i e_i$, 今要证 $s_n \to x$. 这是定理证明的核心部分. 首先不难用式(1.5.3),式(1.5.4)直接计算得 $\langle s_n, s_n - x\rangle = 0$, 即 $s_n \perp (s_n - x)$. 于是由式(1.5.11),式(1.5.12)有

$$\begin{cases} \| x \|^2 = \| s_n \|^2 + \| x - s_n \|^2, & (1.5.16\text{a}) \\ \| s_n \|^2 = \sum_{i=1}^{n} |\hat{x}_i|^2, & (1.5.16\text{b}) \\ \| s_m - s_n \|^2 = \sum_{n < i \leq m} |\hat{x}_i|^2. & (1.5.16\text{c}) \end{cases}$$

① 对于有限或无限不可数标准正交系,亦可建立类似结论,而且所用方法与这里的方法并无本质差别.

由式(1.5.16a),式(1.5.16b)推出

$$\sum_{i=1}^{\infty}|\hat{x}_i|^2=\lim_n\|s_n\|^2\leqslant\|x\|^2,$$

可见级数 $\sum|\hat{x}_i|^2$ 收敛.这结合式(1.5.16c)得出 $\{s_n\}$ 为 Cauchy 列.设 $s_n\to y\in H(n\to\infty)$.只要证 $x=y$;因可用条件(iii),故又只要证 $\langle y-x,e_i\rangle=0$,即 $\hat{y}_i=\hat{x}_i(\forall i\in\mathbf{N})$.取定 $i\in\mathbf{N}$,有

$$\hat{y}_i=\lim_n\langle s_n,e_i\rangle=\lim_n\langle\sum_{j=1}^{n}\hat{x}_je_j,e_i\rangle=\hat{x}_i.\quad(\text{用式(1.5.6)})$$

(i)⇔(iv).式(1.5.16a)表明 $s_n\to x\Leftrightarrow\|s_n\|^2\to\|x\|^2$,这正说明(i)⇔(iv).

直接看出(i)⇒(v)⇒(iv),故定理得证. □

若 $\{e_i:i\in\mathbf{N}\}$ 是 H 的标准正交基,记 $\hat{x}_i=\langle x,e_i\rangle$,则依式(1.5.14),式(1.5.15)有

$$\|x\|^2=\sum_i|\hat{x}_i|^2,\tag{1.5.14$'$}$$

$$\langle x,y\rangle=\sum_i\hat{x}_i\bar{\hat{y}}_i.\tag{1.5.15$'$}$$

这就表明,H 中的范数与内积都可通过正交坐标来表示,而且其表达式正如 l^2 中的范数与内积公式(见式(1.5.8))一样.实际上,映射

$$T:H\to l^2,\quad x\to(\hat{x}_i)$$

正是一等距同构.由此可见,H 与 l^2 作为 Hilbert 空间实质上并无不同.这就完全阐明了 Hilbert 空间 H 的结构.以上所述同时也表明了,对于 Hilbert 空间问题的研究,标准正交基有明显的优势.首先,正交坐标有极简单的计算公式:$\hat{x}_i=\langle x,e_i\rangle$;而且,利用正交坐标计算范数的公式也足够规范而便于运用.这些优势是 Schauder 基或基本集等根本无法比拟的.正是这类差别使得 Hilbert 空间方法比 Banach 空间方法要简便得多.

当然,以上结论的前提是某个标准正交基 $\{e_i\}$ 存在.然而,必定有这样的基存在吗?回答是肯定的,而且有求出标准正交基的普遍方法.设 H 是一个可分的无限维 Hilbert 空间[①],取 H 的可数基本集 $\{x_i\}$(参看§1.3.3),可设 $\{x_i\}$ 线性无关(否则可除去一些多余的元素),然后依如下的 **Schimidt 正交化方法**将其标准正交

① 有限维内积空间的标准正交基已在线性代数中得到处理,此处不必重复;至于不可分的 Hilbert 空间,原则上亦有确定其标准正交基的类似方法,但这种情况在应用上并不具普遍性,不予考虑.

化.令

$$\begin{cases} e_1 = x_1 / \| x_1 \|, \\ e_i = \dfrac{x_i - \sum_{j<i} \langle x_i, e_j \rangle e_j}{\left\| x_i - \sum_{j<i} \langle x_i, e_j \rangle e_j \right\|}, \quad i = 2,3,\cdots, \end{cases} \tag{1.5.17}$$

则 $\{e_i : i \in \mathbf{N}\}$ 是一标准正交系；它如 $\{x_i\}$ 一样亦为 H 的基本集，因此依定理 1.5.5是 H 的标准正交基.

§1.5.3　标准正交基的例子

现在考虑空间 $L^2(J)$(J 是某个实区间)的一些常用的标准正交基，它们通常由多项式或三角函数组成.

(i) **三角函数系**.取 $J = [-\pi, \pi]$. 如同多项式之全体在 $L^2(J)$ 中稠密一样，三角多项式之全体亦在 $L^2(J)$ 中稠密①，形如

$$\text{const} + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx)$$

的函数称为三角多项式.这就表明，函数系

$$\frac{1}{\sqrt{2\pi}}, \quad \frac{\cos nx}{\sqrt{\pi}}, \quad \frac{\sin nx}{\sqrt{\pi}}, \quad n = 1,2,\cdots \tag{1.5.18}$$

是 $L^2(J)$ 的基本集.另一方面，在微积分学中已确立函数系(1.5.18)式在 J 上的正交性，且易验知其中每个函数有 L^2 范数 1.因此式(1.5.18)是 $L^2(J)$ 的标准正交基.于是，每个 $u \in L^2(J)$ 可展开为均方收敛(注意这是要点)的 Fourier 级数：

$$u(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx), \tag{1.5.19}$$

其中 a_n, b_n 是通常的 Fourier 系数.值得注意的是，并不能由定理 1.5.5 推出，展开式(1.5.19)在通常收敛的意义下成立.不过，确实有一个出色的定理断言，展开式(1.5.19)在 J 上几乎处处成立.

若将 $L^2(J)$ 中的内积定义修改为

$$\langle u, v \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} u(x) \overline{v(x)} \mathrm{d}x, \tag{1.5.20}$$

这相当于以测度 $\pi^{-1}\mathrm{d}x$ 取代 Lebesgue 测度 $\mathrm{d}x$，则空间 $L^2(J, \pi^{-1}\mathrm{d}x)$ 有标准正交

① 这是实分析的一个标准结果.

基：

$$\frac{1}{\sqrt{2}}, \quad \cos nx, \quad \sin nx, \quad n = 1,2,\cdots. \tag{1.5.18$'$}$$

设 a_n, b_n 依式(1.5.19)，则依内积式(1.5.20)有

$$a_0 = \sqrt{2}\left\langle u, \frac{1}{\sqrt{2}}\right\rangle, \quad a_n = \langle u, \cos nx\rangle, b_n = \langle u, \sin nx\rangle \quad (n \geqslant 1).$$

于是由式(1.5.14)得出通常意义下的 Parseval 等式：

$$\frac{1}{\pi}\int_{-\pi}^{\pi} |u(x)|^2 \mathrm{d}x = \frac{|a_0|^2}{2} + \sum_{n=1}^{\infty}(|a_n|^2 + |b_n|^2).$$

一般 Hilbert 空间中依标准正交基的展开，在很大程度上受到 Fourier 展开的启发．因此，也称展开式(1.5.13)为 x 关于标准正交基 $\{e_i\}$ 的 Fourier 级数，而称其系数 $\alpha_i = \langle x, e_i\rangle$ 为 x 关于基 $\{e_i\}$ 的 **Fourier 系数**.

(ii) **Legendre 多项式系**．取 $J = [-1,1]$．如例 1.3.8(iii)已指明的，$L^2(J)$ 有基本集 $\{1, x, x^2, \cdots\}$；将其标准正交化，得到一多项式系 $\{L_n : n \geqslant 0\}$，称为 Legendre 多项式系．由定理 1.5.5，$\{L_n\}$ 是 $L^2(J)$ 的标准正交基．从稍不同的途径，能得出 L_n 的如下通式(参考[1])

$$L_n(x) = \frac{1}{2^n n!}\sqrt{\frac{2n+1}{2}}\frac{\mathrm{d}^n}{\mathrm{d}x^n}(x^2-1)^n \quad (n \geqslant 0), \tag{1.5.21}$$

其中开头几项是

$$L_0 = \frac{1}{\sqrt{2}}, \quad L_1 = \sqrt{\frac{3}{2}}x, \quad L_2 = \sqrt{\frac{5}{8}}(3x^2-1), \cdots \tag{1.5.21$'$}$$

若取 $J = [a,b](a < b)$，则通过线性代换

$$\varphi(x) = \frac{2x-a-b}{b-a},$$

可利用 $\{L_n\}$ 构成空间 $L^2[a,b]$ 的如下标准正交基

$$\sqrt{\frac{2}{b-a}}L_n(\varphi(x)) = \sqrt{\frac{2}{b-a}}L_n\left(\frac{2x-a-b}{b-a}\right) \quad (n \geqslant 0). \tag{1.5.22}$$

(iii) **Hermite 多项式系**．可验证 $\{x^n \mathrm{e}^{-x^2/2} : n \geqslant 0\}$ 是空间 $L^2(\mathbf{R})$ 的基本集．显

然,其中的 $e^{-x^2/2}$ 起衰减因子的作用,它保证 $x^n e^{-x^2/2}$ 在 $\mathbf{R}$ 上平方可积.将上述函数系标准正交化,得到 $L^2(\mathbf{R})$ 的一个标准正交基 $\{e_n : n \geqslant 0\}$. 利用公式(1.5.17)不难归纳地说明, e_n 是 $e^{-x^2/2}$ 与一个 n 次多项式的乘积,将它表成

$$e_n(x) = (2^n n!\sqrt{\pi})^{-1/2} e^{-x^2/2} H_n(x) \quad (n \geqslant 0),$$

其中的 n 次多项式 H_n 即称为 Hermite 多项式. H_n 是 2 阶齐次线性微分方程

$$u'' - 2xu' + 2nu = 0$$

的解,它有如下通式(参看文献[1]第 172 页):

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2} \quad (n \geqslant 0), \tag{1.5.23}$$

其中开头几项是

$$H_0 = 1, \quad H_1 = 2x, \quad H_2 = 4x^2 - 2, \cdots. \tag{1.5.23$'$}$$

若将 $L^2(\mathbf{R})$ 中的内积定义修改为

$$\langle u, v \rangle = \int_{-\infty}^{\infty} u(x)\overline{v(x)} e^{-x^2} dx,$$

这相当于以测度 $d\mu = e^{-x^2} dx$ 取代 Lebesgue 测度 dx, 则空间 $L^2(\mathbf{R}, \mu)$ 以多项式系

$$(2^n n!\sqrt{\pi})^{-1/2} H_n(x), \quad n = 0, 1, 2, \cdots$$

为其标准正交基.

(iv) **Laguerre 多项式系**. 可验证 $\{x^n e^{-x/2} : n \geqslant 0\}$ 是空间 $L^2(\mathbf{R}_+)$ 的基本集. 将其标准正交化,得到 $L^2(\mathbf{R}_+)$ 的一个标准正交基

$$e_n(x) = e^{-x/2} P_n(x) \quad (n \geqslant 0),$$

其中 $P_n(x)$ 是一个 n 次多项式,称为 Laguerre 多项式. P_n 是如下微分方程

$$xu'' + (1 - x)u' + nu = 0$$

的解,它有如下通式(参看文献[1]第 175 页):

$$P_n(x) = \frac{e^x}{n!} \frac{d^n}{dx^n}(x^n e^{-x}) \quad (n \geqslant 0), \tag{1.5.24}$$

其开头几项为

$$P_0 = 1, \quad P_1 = 1 - x, \quad P_2 = 1 - 2x + \frac{x^2}{2}, \cdots \tag{1.5.24$'$}$$

注意，$\{P_n\}$ 正是空间 $L^2(\mathbf{R}_+, \mathrm{e}^{-x}\mathrm{d}x)$ 的标准正交基.

(v) **Haar 函数系**. 以 φ_{nk} 记区间 $\left(\frac{k-1}{2^n}, \frac{k}{2^n}\right)$ 的特征函数，令

$$h_{nk} = c_n(\varphi_{n,2k-1} - \varphi_{n,2k}) \quad (k = 1,2,\cdots,2^{n-1}, n = 1,2,\cdots).$$

若 $k \neq l$，则 $h_{nk}h_{nl} = 0$；若 $m < n$，则

$$\int_0^1 h_{nk}(x)h_{ml}(x)\mathrm{d}x = \text{const}\int_{(2k-2)/2^n}^{2k/2^n} h_{nk}(x)\mathrm{d}x = 0,$$

可见 $\{h_{nk}\}$ 是 $J = [0,1]$ 上的正交系. 其次，

$$\int_0^1 |h_{nk}(x)|^2\mathrm{d}x = \frac{c_n^2}{2^{n-1}},$$

可见只要取 $c_n = \sqrt{2^{n-1}}$，就可使 $\{h_{nk}\}$ 为空间 $L^2(J)$ 中的标准正交系. 显然，在其中补充函数 $h_{00} \equiv 1$ 之后仍为标准正交系. 今指明 $\{h_{nk}\}$ 为空间 $L^2(J)$ 的标准正交基. 由定理 1.5.5，只要说明 $\{h_{nk}\}$ 是极大正交系. 设 $u \in \{h_{nk}\}^{\perp}$，则从

$$\langle u, c_n^{-1}h_{nk} \pm c_{n+1}^{-1}h_{n+1,2k-1}\rangle = 0$$

推出

$$\int_{\frac{2k-2}{2^n}}^{\frac{2k-1}{2^n}} u(x)\mathrm{d}x = \int_{\frac{2k-1}{2^n}}^{\frac{2k}{2^n}} u(x)\mathrm{d}x = 0,$$

从而 $F(x) = \int_0^x u(y)\mathrm{d}y$ 在点 $x = k2^{-n}(k = 0,1,\cdots,2^n, n = 1,2,\cdots)$ 为零. 这推出 $F(x) \equiv 0$，因此 $u(x) = 0$, a.e..

如上的函数系称为 Haar 函数系，它与前面所讨论的函数系的明显不同之处是，其中的函数(h_{00} 除外)是跳跃的. 这一特点使它在信号处理方面多有用处.

§1.5.4 最佳逼近

如上节中已考虑过的(参看推论 1.4.5)，最佳逼近问题可描述为：对于给定的集 $A \subset H$ 与点 $x \in H$，求一点 $a \in A$，使得 $\| x - a \| = d(x, A)$；这意味着 a 是最小化问题

$$\min \| x - a \|, \quad a \in A \tag{1.5.25}$$

的最优解. 若 A 是 H 的子空间，则来自 Euclid 几何的一个启示是：问题(1.5.25)式

的解 a 应是从点 x 向 A 所引垂线的垂足.正是这一提示导向内积空间中最佳逼近问题的完全解决.

定理 1.5.6　设 A 是 H 的完备子空间，$x \in H$.

(i) 存在惟一性：x 在 A 中有惟一最佳逼近 a，且 $x - a \in A^{\perp}$.

(ii) 用 A 的基求最佳逼近：若 $\{a_1, a_2, \cdots, a_n\}$ 是 A 的基，a 是 x 在 A 中的最佳逼近，则 $a = \sum_i \beta_i a_i$，

$$\beta = (\beta_1, \beta_2, \cdots, \beta_n)^{\mathrm{T}} = G^{-1}(\langle x, a_1 \rangle, \cdots, \langle x, a_n \rangle)^{\mathrm{T}}, \qquad (1.5.26)$$

其中 $G = [\langle a_j, a_i \rangle]_{n \times n}$ 称为向量组 $\{a_i\}$ 的 **Gram 矩阵**.

(iii) 用 A 的标准正交基求最佳逼近：若 $\{e_1, e_2, \cdots, e_n\}$ 是 A 的标准正交基，a 是 x 在 A 中的最佳逼近，则

$$a = \sum_i \langle x, e_i \rangle e_i. \qquad (1.5.27)$$

证　(i) 取序列 $\{x_n\} \subset A$，使得 $\| x - x_n \| \to \rho \triangleq d(x, A)(n \to \infty)$. 分别以 $x_m - x$ 与 $x_n - x$ 替换 x 与 y 应用中线公式(1.5.9)，得

$$\begin{aligned} \| x_m - x_n \|^2 &= \| (x_m - x) - (x_n - x) \|^2 \\ &= 2\| x_m - x \|^2 + 2\| x_n - x \|^2 - 4\left\| \frac{x_m + x_n}{2} - x \right\|^2 \\ &\leqslant 2\| x_m - x \|^2 + 2\| x_n - x \|^2 - 4\rho^2 \to 0, \quad (m, n \to \infty) \end{aligned}$$

可见 $\{x_n\}$ 是 Cauchy 列.因 A 完备，故可设 $x_n \to a \in A(n \to \infty)$. 显然 $\| x - a \| = \rho$. 若另有 $a' \in A$ 使 $\| x - a' \| = \rho$，则再用中线公式得

$$\begin{aligned} \| a - a' \|^2 &= 2\| a - x \|^2 + 2\| a' - x \|^2 - 4\left\| \frac{a + a'}{2} - x \right\|^2 \\ &\leqslant 2\rho^2 + 2\rho^2 - 4\rho^2 = 0, \end{aligned}$$

可见 $a = a'$. 因此，a 是 x 在 A 中的惟一最佳逼近.

令 $b = x - a$，今证 $b \in A^{\perp}$. 取定 $y \in A$，$\forall \lambda \in \mathbf{K}$，有

$$\begin{aligned} \rho^2 &\leqslant \| x - (a + \lambda y) \|^2 = \| b - \lambda y \|^2 \\ &= \rho^2 - \lambda \langle y, b \rangle - \lambda \langle b, y \rangle + |\lambda|^2 \| y \|^2. \end{aligned}$$

可设 $y \neq 0$ (否则已不必证). 以 $\lambda = \langle b, y \rangle / \| y \|^2$ 代入以上不等式，整理后得 $|\langle b, y \rangle|^2 \leqslant 0$，故 $\langle b, y \rangle = 0$. $b \in A^{\perp}$ 得证.

(ii) 设 $a = \sum \beta_i a_i$ 是 x 在 A 中的最佳逼近，则由 $x - a \in A^{\perp}$ 有 $\langle x - a, a_i \rangle$

$= 0$, 即

$$\langle x, a_i \rangle - \langle a, a_i \rangle = \sum_j \beta_j \langle a_j, a_i \rangle \quad (1 \leqslant i \leqslant n),$$

或

$$G\beta = (\langle x, a_1 \rangle, \cdots, \langle x, a_n \rangle)^{\mathrm{T}}. \tag{1.5.28}$$

用标准的线性代数方法易从 $\{a_i\}$ 线性无关推出 Gram 矩阵 G 可逆. 于是可从线性方程(1.5.28)解出 β, 从而得出式(1.5.26).

(iii) 关于标准正交基的 Gram 矩阵显然为单位矩阵,因此直接从式(1.5.26)推出式(1.5.27). □

若取 $H = L^2$, A 为 L^2 的 n 维子空间,则定理 1.5.6 提供了求任何 $u \in L^2$ 在 A 中的最佳均方逼近的简便方法. 这一方法应用于各个领域,试看下面几个例子.

例 1.5.7(最佳均方逼近多项式) 在 $J = [0,1]$ 上求最佳均方逼近 $u(x) = \mathrm{e}^x$ 的 2 次多项式.

解法 I(应用公式(1.5.26)) 以 A 记 J 上次数 $\leqslant 2$ 的多项式之全体,则 $\{1, x, x^2\}$ 是 A 的基,其 Gram 矩阵为

$$G = \begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{bmatrix}.$$

以 $\langle \cdot, \cdot \rangle$ 记 $L^2[0,1]$ 中的内积,依次算出

$$\begin{cases} \langle u, 1 \rangle = \int_0^1 \mathrm{e}^x \mathrm{d}x = \mathrm{e} - 1, \\ \langle u, x \rangle = \int_0^1 x\mathrm{e}^x \mathrm{d}x = 1, \\ \langle u, x^2 \rangle = \int_0^1 x^2 \mathrm{e}^x \mathrm{d}x = \mathrm{e} - 2. \end{cases} \tag{1.5.29}$$

于是,依公式(1.5.26)得 $u = \mathrm{e}^x$ 的最佳均方逼近 2 次多项式为

$$\begin{aligned} P(x) &= (1, x, x^2) G^{-1} (\mathrm{e} - 1, 1, \mathrm{e} - 2)^{\mathrm{T}} \\ &= 39\mathrm{e} - 105 + (588 - 216\mathrm{e})x + (210\mathrm{e} - 570)x^2 \\ &\approx 1.0130 + 0.8511x + 0.8392x^2. \end{aligned}$$

解法 II(应用公式(1.5.27)) 结合式(1.5.21)′与式(1.5.22),知

$$\varphi_0 = 1,\quad \varphi_1 = \sqrt{3}(2x-1),\quad \varphi_2 = \sqrt{5}(6x^2-6x+1)$$

是 A 的标准正交基. 依次算出(注意用式(1.5.29)简化计算)

$$\langle u,\varphi_0\rangle = \int_0^1 e^x dx = e-1;$$

$$\langle u,\varphi_1\rangle = \sqrt{3}\int_0^1 (2x-1)e^x dx = \sqrt{3}(3-e);$$

$$\langle u,\varphi_2\rangle = \sqrt{5}\int_0^1 (6x^2-6x+1)e^x dx = \sqrt{5}(7e-19).$$

于是依公式(1.5.27)有

$$\begin{aligned} P(x) &= e-1+3(3-e)(2x-1)+5(7e-19)(6x^2-6x+1) \\ &= 39e-105+(588-216e)x+(210e-570)x^2, \end{aligned}$$

这恰如解法Ⅰ的结果. □

例 1.5.8(方差逼近)　设 Y 与 $X_1,X_2,\cdots,X_n$ 是概率空间 Ω 上的 2 阶随机变量,假定 $\{X_i\}$ 线性无关,求 $X = \sum\beta_i X_i$,使得

$$E|Y-\sum_i \beta_i X_i|^2 = \min,$$

$E(\cdot)$ 表期望. 如上的 X 称为 Y 的**方差逼近**或均方逼近.

解　$L^2(\Omega)$ 依内积 $\langle X,Y\rangle = E(X\overline{Y})$ 是一个 Hilbert 空间,所求的 $X = \sum\beta_i X_i$ 正是 Y 在子空间 $A \triangleq \mathrm{span}\{X_1,X_2,\cdots,X_n\}$ 中的最佳逼近. 于是,直接用公式(1.5.26)有

$$\beta \triangleq (\beta_1,\beta_2,\cdots,\beta_n)^T = G^{-1}(E(Y\overline{X}_1),\cdots,E(Y\overline{X}_n))^T, \qquad (1.5.30)$$

其中 $G = [E(X_j\overline{X}_i)]_{n\times n}$. □

例 1.5.8 的一个典型应用是平稳随机序列的**线性外推问题**. 设 $\{X_t : t\in \mathbf{Z}\}$ 是一**平稳随机序列**(此类序列常用于描述某些信号过程),今要以某个

$$\widetilde{X}_{t+m} = \sum_{i=1}^{n}\beta_i X_{t+i}\quad (m > n \geqslant 1),$$

作为 X_{t+m} 的预测,使得均方误差最小,即

$$E|X_{t+m} - \widetilde{X}_{t+m}|^2 = \min.$$

设 $R(t)$ 是序列 $\{X_t\}$ 的**相关函数**,这意味着 $E(X_s\overline{X}_t) = R(s-t)$,则依公式(1.5.30)有

$$(\beta_1,\beta_2,\cdots,\beta_n)^{\mathrm T}=G^{-1}[R(m-1),\cdots,R(m-n)]^{\mathrm T},$$

其中 $G=[R(j-i)]_{n\times n}$.

例 1.5.9(最小二乘法) 设对于随机变量 $Y,X_j(1\leqslant j\leqslant m)$ 取得 n 组样本值 $y_i,x_{ij}(1\leqslant i\leqslant n,1\leqslant j\leqslant m),n>m$,求估计式 $Y=\sum\beta_jX_j$,使得

$$\sum_i\left(y_i-\sum_j\beta_jx_{ij}\right)^2=\min.$$

解 若令 $y=(y_1,y_2,\cdots,y_n)^{\mathrm T},x_j=(x_{1j},x_{2j},\cdots,x_{nj})^{\mathrm T}$,则问题相当于求 $a=\sum\beta_jx_j\in A\triangleq\operatorname{span}\{x_1,x_2,\cdots,x_m\}$,使得 a 是 $y\in\mathbf R^n$ 在 A 中的最佳逼近.不妨设 $\{x_1,x_2,\cdots,x_m\}$ 线性无关,这意味着矩阵 $X=[x_{ij}]$ 的秩为 m. 令 $\beta=(\beta_1,\beta_2,\cdots,\beta_m)^{\mathrm T}$,则依式(1.5.26)有

$$\beta=\begin{bmatrix}x_1^{\mathrm T}x_1 & x_1^{\mathrm T}x_2 & \cdots & x_1^{\mathrm T}x_m\\ x_2^{\mathrm T}x_1 & x_2^{\mathrm T}x_2 & \cdots & x_2^{\mathrm T}x_m\\ \vdots & \vdots & & \vdots\\ x_m^{\mathrm T}x_1 & x_m^{\mathrm T}x_2 & \cdots & x_m^{\mathrm T}x_m\end{bmatrix}^{-1}\begin{bmatrix}x_1^{\mathrm T}y\\ x_2^{\mathrm T}y\\ \vdots\\ x_m^{\mathrm T}y\end{bmatrix}\tag{1.5.31}$$

$$=(X^{\mathrm T}X)^{-1}X^{\mathrm T}y,$$

这正是回归分析中熟知的**最小二乘估计**公式.

定理 1.5.6 的证明实际上已达到以下结论.

定理 1.5.10(正交分解定理) 设 A 是 Hilbert 空间 H 的闭子空间,则有拓扑直和分解 $H=A\oplus A^{\perp}$. 设 P_A 是由这一分解决定的投影,则 $P_Ax(x\in H)$ 是 x 在 A 中的最佳逼近.

证 任给 $x\in H$,以 a 记 x 在 A 中的最佳逼近,则由定理 1.5.6 有 $x=a+b,b\in A^{\perp}$. 故得 $H=A+A^{\perp}$. 若 $x\in A\cap A^{\perp}$,则必 $0=\langle x,x\rangle=\|x\|^2$,因而 $x=0$,故得 $A\cap A^{\perp}=\{0\}$. 因此 $H=A\oplus A^{\perp}$. 显然 $A^{\perp}$ 亦为 H 的闭子空间,故 $A\oplus A^{\perp}$ 是拓扑直和.余下的结论是明显的. □

定理 1.5.10 中的分解 $H=A\oplus A^{\perp}$ 称为**正交分解**. P_A 称为从 H 到 A 的**正投影**,它由 A 惟一决定.

由正交分解定理直接推出,若 A 是 H 的闭子空间,则 A 与 $A^{\perp}$ 互为正交补,即 $A=A^{\perp\perp}$. 今利用这一点来解以下问题.

例 1.5.11(最小范数问题) 给定线性无关向量组 $\{a_1,a_2,\cdots,a_n\}\subset H$ 与 $\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_n)^{\mathrm T}\in\mathbf K^n$,求解约束最小化问题:

$$\min \| x \|, \quad \text{s.t.} \langle x, a_i \rangle = \alpha_i (1 \leqslant i \leqslant n). \tag{1.5.32}$$

解 表面上看,此问题似乎与最佳逼近问题无关,但我们可将它转化为一个最佳逼近问题.关键在于取 $x_0 \in H$,使得

$$\langle x_0, a_i \rangle = \alpha_i \quad (1 \leqslant i \leqslant n).$$

令 $x = x_0 - y$,则 x 满足式(1.5.32)中的约束条件$\Leftrightarrow \langle y, a_i \rangle = 0 (1 \leqslant i \leqslant n)$,于是问题(1.5.32)等价于

$$\min \| x_0 - y \|, \quad y \in A^{\perp}, \tag{1.5.32$'$}$$

其中 $A = \text{span}\{a_1, a_2, \cdots, a_n\}$. 由定理 1.5.6,式(1.5.32)′的解 y 是x_0 在 $A^{\perp}$ 上的正投影.于是问题(1.5.32)的解 $x = x_0 - y$ 是x_0 在 $A = A^{\perp\perp}$ 上的正投影.由定理 1.5.6(ii),有 $x = \sum \beta_i a_i$,

$$\beta = (\beta_1, \beta_2, \cdots, \beta_n)^{\mathrm{T}} = G^{-1} \alpha, \quad (\text{用式}(1.5.26))$$

其中 G 是向量组 $\{a_i\}$ 的 Gram 矩阵. □

图 1-8 给出了例 1.5.11 的几何解释,其中 $x_0 + A^{\perp}$ 表示由约束条件 $\langle x, a_i \rangle = \alpha_i (1 \leqslant i \leqslant n)$ 决定的线性流形.

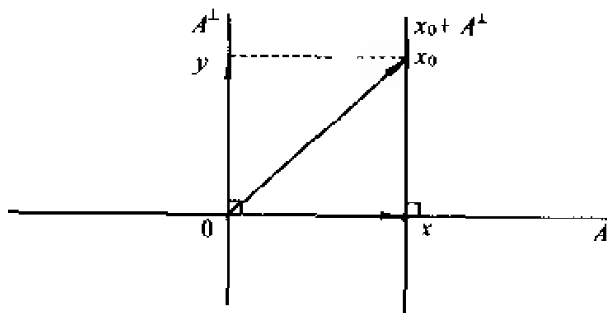


图 1-8

§ 1.6 Sobolev 空间

在 § 1.2 中讨论的空间 $L^p(\Omega)$ 与 $C^m(\Omega)$,虽然都属于应用上最重要的函数空间,但两者各有一些很明显的缺点.空间 $L^p(\Omega)$ 的主要缺点是,它完全不考虑函数的导数.即使 $u \in L^p(\Omega)$ 可微,u 的导数在L^p 范数 $\| u \|_p$ 的定义中也完全不起作用,因而距离 $\| u - v \|_p$ 完全不能反映u 与v 的导数之间的差别.因此,在涉及微分的问题(如微分方程理论)中,空间 L^p 难以充分发挥作用.空间 $C^m(\Omega)$ 虽

然没有 $L^p(\Omega)$ 的上述缺点,但其空间结构却不能令人满意.在下章中我们将看到,空间 $L^p(1 < p < \infty)$ 的自反性是一个导致许多深刻结论的良好性质.而 $C^m(\Omega)$ 却不具有自反性,更不像 L^2 那样是 Hilbert 空间.概括起来,似乎可以说,空间 $L^p(\Omega)$ 包含坏的函数却有好的空间结构;而与此相反,空间 $C^m(\Omega)$ 由好函数组成但具有坏的空间结构.如果一种空间能兼有 $L^p(\Omega)$ 与 $C^m(\Omega)$ 的优点,就必将避免两者的缺点.这样一种理想的函数空间,就是本节要介绍的 Sobolev 空间,它是俄罗斯数学家 Sobolev 在研究数学物理问题时首先引入的.

与空间 $L^p(\Omega)$,$C^m(\Omega)$ 相比,Sobolev 空间无疑具有更复杂的构成,关于它的完全表述要用到较深入的知识(如广义函数、Fourier 分析等),因而无法在目前准备欠缺的情况下详细介绍.在本节中,我们仅对 Sobolev 空间的主要特征作简要的勾画,而避开可能需要更多预备知识的论证细节.欲知其详的读者可参考文献[3].

以下给定 $p \in [1,\infty)$,$m \in \mathbf{Z}_+$,开区域 $\Omega \subset \mathbf{R}^n$.约定 α 记 n 重指标 $(\alpha_1, \alpha_2, \cdots, \alpha_n) \in \mathbf{Z}_+^n$,$|\alpha| = \sum \alpha_i$.$\partial^\alpha u$ 依(1.2.12)式.

§1.6.1 空间 $H^{m,p}(\Omega)$

在§1.3.3 中已经指明,每个 $u \in L^p(\Omega)$ 都可用连续函数 L^p 逼近(参考例 1.3.8(v)).这就意味着,若令

$$X = \{u \in C(\Omega) : \|u\|_p < \infty\},$$

则 X 是 $L^p(\Omega)$ 的稠密子空间,因而可将 $L^p(\Omega)$ 看作赋范空间 $(X, \|\cdot\|_p)$ 的完备化.将上述思想加以引伸,可用来构成一类新的空间,它有一些类似于空间 $L^p(\Omega)$ 的特点.

定义 1.6.1 对任给 $u \in C^m(\Omega)$,规定

$$\|u\|_{m,p} = \Big(\sum_{|\alpha| \leqslant m} \|\partial^\alpha u\|_p^p\Big)^{1/p}. \tag{1.6.1}$$

令

$$S = \{u \in C^m(\Omega) : \|u\|_{m,p} < \infty\}, \tag{1.6.2}$$

则 $(S, \|\cdot\|_{m,p})$ 是一个赋范空间(试验证),因而其完备化是一个 Banach 空间,称之为 **Sobolev 空间**,记作 $H^{m,p}(\Omega)$,其中的范数仍记作 $\|\cdot\|_{m,p}$.

特别,若取 $p = 2$,则 S 是一个内积空间,其中的内积定义为

$$\langle u, v\rangle_m = \sum_{|\alpha| \leqslant m} \langle \partial^\alpha u, \partial^\alpha v\rangle \quad (u, v \in S), \tag{1.6.3}$$

其中 $\langle\cdot,\cdot\rangle$ 记 $L^2(\Omega)$ 中的内积.于是,$H^{m,2}(\Omega)$ 是一个 Hilbert 空间,通常将它简写作 $H^m(\Omega)$.

你已看到,利用完备化来定义我们所需要的新空间 $H^{m,p}(\Omega)$,在逻辑上是一件很简单的事.但在直观上却留下一个遗憾:我们不知道该如何看待 $H^{m,p}(\Omega)$ 中附加进来(即 S 之外)的元素,它们是 Ω 上的函数吗? 若 $u \in H^{m,p}(\Omega) \setminus S$,该如何计算 $\| u \|_{m,p}$? 幸而有以下结果.

命题 1.6.2 对任给 $u \in H^{m,p}(\Omega)$,存在惟一函数组 $\{u^\alpha : |\alpha| \leqslant m\} \subset L^p(\Omega)$,使得若 $\{u_k\} \subset S$,在 $H^{m,p}(\Omega)$ 中 $u_k \to u(k \to \infty)$,则

$$\partial^\alpha u_k \xrightarrow{L^p} u^\alpha \quad (k \to \infty, |\alpha| \leqslant m). \tag{1.6.4}$$

证 取定 $u \in H^{m,p}(\Omega)$. 由 $H^{m,p}(\Omega)$ 的定义,有序列 $\{u_k\} \subset S$,使得 $u_k \to u(k \to \infty)$. $\{u_k\}$ 依范数 $\| \cdot \|_{m,p}$ 必为 Cauchy 列;于是由式(1.6.1)推出,序列 $\{\partial^\alpha u_k\}$ ($|\alpha| \leqslant m$) 均为 $L^p(\Omega)$ 中的 Cauchy 列,因此可设 $\partial^\alpha u_k \xrightarrow{L^p} u^\alpha \in L^p(\Omega)(k \to \infty, |\alpha| \leqslant m)$. 这就得出了满足式(1.6.4)的函数组 $\{u^\alpha\} \subset L^p(\Omega)$.

下面只要证惟一性. 设序列 $\{v_k\} \subset S$ 在 $H^{m,p}(\Omega)$ 中收敛于 u,$\partial^\alpha v_k \xrightarrow{L^p} v^\alpha \in L^p(\Omega)(k \to \infty, |\alpha| \leqslant m)$. 合并 $\{u_k\}$ 与 $\{v_k\}$ 为序列 $\{w_k\}$,则必 $\| w_k - u \|_{m,p} \to 0$,

$$\partial^\alpha w_k \xrightarrow{L^p} u^\alpha \quad \partial^\alpha w_k \xrightarrow{L^p} v^\alpha. \quad (k \to \infty, |\alpha| \leqslant m).$$

这就得出 $u^\alpha = v^\alpha$,a.e.($|\alpha| \leqslant m$),惟一性得证. □

在命题 1.6.2 中,若 $u \in C^m(\Omega)$,则必定 $u^\alpha = \partial^\alpha u$,a.e.($|\alpha| \leqslant m$);若 $u \notin C^m(\Omega)$,则约定称 u^α 为 u 的 α 阶**广义导数**,亦记作 $\partial^\alpha u$($|\alpha| \leqslant m$),它几乎处处由 u 惟一确定. 自然将 u 等同于 $\partial^0 u$ (即 u^0,0 是 $\mathbf{Z}^n$ 中的零元). 这就得到集包含关系:

$$S \subset H^{m,p}(\Omega) \subset L^p(\Omega). \tag{1.6.5}$$

因此可以说,$H^{m,p}(\Omega)$ 中的元素无非是普通的 p 次可积函数. 当 $m = 0$ 时,实际上有 $H^{0,p}(\Omega) = L^p(\Omega)$.

但不应因有式(1.6.5)而误解为 $H^{m,p}(\Omega)$ 是 $L^p(\Omega)$ 的赋范子空间,须知在 $H^{m,p}(\Omega)$ 中采用范数 $\| \cdot \|_{m,p}$ 而非 L^p 范数. 设 u, u_k 依命题 1.6.2,则

$$\| u \|_{m,p} = \lim_k \| u_k \|_{m,p} = \lim_k \Big(\sum_{|\alpha| \leqslant m} \| \partial^\alpha u_k \|_p^p \Big)^{1/p}$$

$$= \Big(\sum_{|\alpha| \leqslant m} \| u^\alpha \|_p^p \Big)^{1/p} = \Big(\sum_{|\alpha| \leqslant m} \| \partial^\alpha u \|_p^p \Big)^{1/p}.$$

这就表明,范数公式(1.6.1)亦适用于任何 $u \in H^{m,p}(\Omega)$,只是其中的 $\partial^\alpha u$($|\alpha| \leqslant$

m) 应理解为广义导数. 由此又推出, 对任给 $\{u_k\} \subset H^{m,p}(\Omega)$, 有

$$\| u_k \|_{m,p} \to 0 \Leftrightarrow \partial^\alpha u_k \xrightarrow{L^p} 0 \quad (k \to \infty, |\alpha| \leqslant m).$$

这就表明, $H^{m,p}(\Omega)$ 中的序列收敛归结为相应的导数序列(直到 m 阶为止)的 L^p 收敛. 可以说, 这正好结合了空间 $C^m(\Omega)$ 与 $L^p(\Omega)$ 两者的特点.

Sobolev 空间 $H^{m,p}(\Omega)$ 的一个主要优点是, 它如同 $C^m(\Omega)$ 一样形成一个随 m 递变的阶梯, m 愈高, $H^{m,p}(\Omega)$ 中的函数就愈好(这意味着光滑性愈高), 而这一事实在微分方程理论等领域中是至关重要的. 准确的结论由著名的 Sobolev 嵌入定理给出.

定理 1.6.3(Sobolev 嵌入定理) 设 $\Omega \subset \mathbf{R}^n$ 是边界足够光滑的有界区域, $m > \frac{n}{2} + k, k \in \mathbf{Z}_+$, 则有连续的嵌入 $H^m(\Omega) \subset C^k(\bar{\Omega})$. 这意味着: $\forall u \in H^m(\Omega), \exists v \in C^k(\bar{\Omega})$, 使得 $u = v$, a.e., 且 $\| u \|_k \leqslant \text{const} \| u \|_{m,2}$, 其中 $\| \cdot \|_k$ 记空间 $C^k(\bar{\Omega})$ 中的范数(依式(1.2.6)′).

定理 1.6.3 中的 u 与 v 自然不必区别. 于是定理 1.6.3 的结论可表述为: $H^m(\Omega) \subset C^k(\bar{\Omega})$, 且若在 $H^m(\Omega)$ 中 $u_l \to u$, 则

$$\partial^\alpha u_l(x) \rightrightarrows \partial^\alpha u(x) \quad (l \to \infty, x \in \bar{\Omega}, |\alpha| \leqslant k).$$

用一个简单的一维例子来解释嵌入定理.

例 1.6.4 取 $\Omega = (-1,1) \subset \mathbf{R}$, 则依定理 1.6.3 有

$$H^m(\Omega) \subset C^{m-1}(\bar{\Omega}). \quad (m \geqslant 1)$$

(i) 设 $u = \chi_{(0,1)}$, 则 $u \in L^2(\Omega) = H^0(\Omega) \not\subset C(\bar{\Omega})$;

(ii) 设 $u(x) = \max\{x, 0\} (x \in \Omega)$, 则 $u \in H^1(\Omega) \subset C(\bar{\Omega}) \setminus C^1(\bar{\Omega})$;

(iii) 设

$$u(x) = \begin{cases} 0, & -1 < x \leqslant 0, \\ x^2, & 0 < x < 1. \end{cases}$$

则 $u \in C^1(\bar{\Omega})$, 但 $u \overline{\in} C^2(\bar{\Omega})$. 不过, 可如此修改 u 为 $u_k \in C^2(\bar{\Omega})$, 使得在区间 $(-1/k, 1/k)$ 之外 $u_k(x) = u(x)$, 且 u_k, u'_k 与 u''_k 一致有界. 容易看出 $u_k^{(i)} \xrightarrow{L^2} u^{(i)} (i = 0,1,2)$, 因此 $u \in H^2(\Omega)$. 这个例子说明了 $H^2(\Omega) \subset C^1(\bar{\Omega})$ 不能改进为 $H^2(\Omega) \subset C^2(\bar{\Omega})$.

§1.6.2 空间 $H_0^m(\Omega)$

令 $D(\Omega) = C_c^\infty(\Omega)$ (参看式(1.3.10)), 它在广义函数理论中被称为**检验函**

数空间. 每个 $u \in D(\Omega)$ 都可自然地认定属于 $D(\mathbf{R}^n)$，只要对 $x \in \mathbf{R}^n \setminus \Omega$ 置 $u(x) = 0$.

对于 $u \in D(\Omega)$ 可无限制地进行微分与积分运算，因而具有极大的方便性；而通过适当的极限过程，上述优势又可过渡到由 $D(\Omega)$ 中函数逼近的函数. 这就自然提出一个问题：$D(\Omega)$ 是否在 $H^{m,p}(\Omega)$ 中稠密？如果 $\Omega = \mathbf{R}^n$，则回答是肯定的；在其他情况下则未必. 这又促使我们去考虑一个较 $H^{m,p}(\Omega)$ 为小的空间 $H_0^{m,p}(\Omega)$，即 $D(\Omega)$ 在空间 $H^{m,p}(\Omega)$ 中的闭包，它作为 $H^{m,p}(\Omega)$ 的闭子空间是一 Banach 空间. 若 $p = 2$，则 $H_0^{m,2}(\Omega)$ 是一 Hilbert 空间，简记为 $H_0^m(\Omega)$. 空间 $H_0^{m,p}(\Omega)$ 与 $H_0^m(\Omega)$ 都称为 Sobolev 空间，它们在许多领域有重要应用.

空间 $H_0^m(\Omega)$ 的特殊价值基于以下重要不等式.

定理 1.6.5(Poincaré 不等式) 若 Ω 有界，则对 $u \in H_0^m(\Omega)$ 成立

$$\| u \|_{m,2} \leqslant \operatorname{const}\Big(\sum_{|\alpha| = m} \| \partial^\alpha u \|_2^2 \Big)^{1/2}. \tag{1.6.6}$$

证 只需对 $u \in D(\Omega)$ 证明式(1.6.6). 不妨设 $\Omega \subset (0,a)^n, a > 0, (0,a)^n$ 是边长为 a 的 n 维方体. 于是

$$\begin{aligned}
\int_\Omega | u(x) |^2 \mathrm{d}x &= \int_\Omega \left| \int_0^{x_1} \frac{\partial}{\partial t} u(t, x_2, \cdots, x_n) \mathrm{d}t \right|^2 \mathrm{d}x \\
&\leqslant a \int_\Omega \mathrm{d}x \int_0^a \left| \frac{\partial u}{\partial x_1} \right|^2 \mathrm{d}x_1 \qquad (\text{用 Hölder 不等式}) \\
&\leqslant a \int_\Omega \mathrm{d}x \int_0^a | \nabla u(x) |^2 \mathrm{d}x_1 \\
&= a^2 \int_\Omega | \nabla u(x) |^2 \mathrm{d}x,
\end{aligned}$$

其中 $\nabla u(x)$ 记 $u(x)$ 的梯度. 将所得不等式相继用于 $\partial^\alpha u(|\alpha| < m)$，经有限步后可得出不等式(1.6.6). □

Poincaré 不等式的意义在于：若 Ω 有界，则对于空间 $H_0^m(\Omega)$

$$\| u \|_{(m)} \triangleq \Big(\sum_{|\alpha| = m} \| \partial^\alpha u \|_2^2 \Big)^{1/2} \quad (u \in H_0^m(\Omega)) \tag{1.6.7}$$

是与 $\| \cdot \|_{m,2}$ 等价的范数，而 $\| u \|_{(m)}$ 的表达式仅用到 m 阶导数，因而形式上较 $\| \cdot \|_{m,2}$ 简单些. 特别，对 $u \in H_0^1(\Omega)$，有

$$\| u \|_{(1)} = \left(\int_\Omega | \nabla u(x) |^2 \mathrm{d}x \right)^{1/2};$$

对于 $u \in H_0^1(a,b)(-\infty < a < b < \infty)$，有

$$\| u \|_{(1)} = \left(\int_a^b | u'(x) |^2 \mathrm{d}x \right)^{1/2}.$$

通常，在 $H_0^m(\Omega)$（Ω 有界）中不加说明地使用范数(1.6.7)，且就写作 $\| \cdot \|_m$. 相应地，使用内积(对照式(1.6.3))

$$\langle u, v \rangle_{(m)} = \sum_{|\alpha| = m} \langle \partial^\alpha u, \partial^\alpha v \rangle \quad (u, v \in H_0^m(\Omega)), \tag{1.6.8}$$

其中 $\langle \cdot, \cdot \rangle$ 记 $L^2(\Omega)$ 中的内积.

§1.6.3 对微分方程问题的应用

Sobolev 空间主要应用于微分方程，特别是偏微分方程问题. 深入的讨论只能是专著的任务. 不过，简单的例子也能说明 Sobolev 空间方法的某些基本思路.

经常用作说明的一个简单例子，就是 Laplace 方程的 Dirichlet 边值问题.

$$\begin{cases} -\Delta u(x) = f(x) \quad (x \in \Omega), \\ u \mid \partial\Omega = 0, \end{cases} \tag{1.6.9}$$

其中 $\Omega \subset \mathbf{R}^n$ 是有界区域，$f \in L^2(\Omega)$ 是已知函数，Δ 是 Laplace 算子，即

$$\Delta u = \sum \partial^2 u / \partial x_i^2.$$

若 $u \in C^2(\bar{\Omega})$ 满足式(1.6.9)(这样的 u 称为边值问题式(1.6.9)的古典解)，则对任给 $v \in D(\Omega)$ 有

$$\begin{aligned} \int_\Omega f v \mathrm{d}x &= \int_\Omega \Delta u \cdot v \mathrm{d}x \\ &= \int_\Omega \nabla u \cdot \nabla v \mathrm{d}x - \int_{\partial\Omega} \frac{\partial u}{\partial n} v \mathrm{d}\sigma \\ &= \int_\Omega \nabla u \cdot \nabla v \mathrm{d}x, \end{aligned}$$

其中用到推广的 Green 公式，$\mathrm{d}\sigma$ 是边界 $\partial\Omega$ 的“面积元”，$\frac{\partial u}{\partial n}$ 是 u 沿边界法向的导数. 因 $D(\Omega)$ 在 $H_0^1(\Omega)$ 中稠密，故所得等式亦应适用于 $v \in H_0^1(\Omega)$，即应有等式

$$\int_\Omega \nabla u \cdot \nabla v \mathrm{d}x = \int_\Omega f v \mathrm{d}x \quad (\forall v \in H_0^1(\Omega)), \tag{1.6.10}$$

或写成

$$\langle u, v \rangle_{(1)} = \langle f, v \rangle \quad (\forall v \in H_0^1(\Omega)), \tag{1.6.10$'$}$$

其中 $\langle\cdot,\cdot\rangle_{(1)}$ 依式(1.6.8).这就启示出一个自然的想法:即使 u 未必是问题(1.6.9)式的古典解,只要它满足式(1.6.10),或许就能起某种广义解的作用.因此,凡满足式(1.6.10)的 $u\in H_0^1(\Omega)$ 称为问题(1.6.9)式的**弱解**.若能判明弱解 $u\in C^2(\Omega)$,则 u 实际上就是古典解.这就将求解(1.6.9)式的问题转化成了寻求其弱解的问题,后者通常要容易些.以下就是关于弱解存在性的一个简单结果.

定理 1.6.6　设 $\Omega\subset\mathbf{R}^n$ 是有界区域,$f\in L^2(\Omega)$,则边值问题(1.6.9)式存在惟一弱解.

定理的证明并不困难,但要用到一些到下章才出现的概念与结论.现在可以粗看一下以下的证明,这将为在下章学习有界线性泛函理论预先作点铺垫.

证　由 $T(v)=\langle f,v\rangle(v\in H_0^1(\Omega))$ 定义出空间 $H_0^1(\Omega)$ 上的一个线性泛函,它满足

$$|T(v)|\leqslant\|f\|_2\|v\|_2\leqslant\mathrm{const}\,\|v\|_{1,2},$$

可见 T 是实 Hilbert 空间 $H_0^1(\Omega)$ 上的有界线性泛函.由 Riesz 表示定理(参看定理2.3.6)存在惟一的 $u\in H_0^1(\Omega)$,使得 $\langle u,v\rangle_1=T(v)(\forall v\in H_0^1(\Omega))$.这正表明 u 满足(1.6.10)′,因此 u 是问题(1.6.9)的惟一弱解.　□

§1.7　其他抽象空间

本章前 6 节给出了 Banach 空间(包括其特款 Hilbert 空间与 Sobolev 空间)基础概念的一个梗概.Banach 空间理论为范围广泛的数学问题提供了一个适当的空间框架.它既不像 Hilbert 空间那样失之过窄,缺少回旋余地;也不像某些更一般的抽象空间(例如拓扑向量空间)那样失之过宽,以至无法充分利用平常 Euclid 空间中的许多有益启示.本书将 Banach 空间放在中心地位,看来是合理的.但应立即指出,在抽象空间理论不断拓广的进程中,Banach 空间只不过是其中的一站而已.更一般的空间概念的引进,不只是逻辑上的自然引伸,更受到多个领域实际需要的强劲推动.对于那些应用上日益重要的抽象空间有所了解,自然成为泛函分析学习中的一项任务.当然,下面的简略介绍仅能使读者见其一斑,但由此所获得的印象,对于希望进入这些空间领域的读者还是有益的.

§1.7.1　度量空间

在§1.1.1 中,已在赋范空间中引进距离:

$$d(x,y)=\|x-y\|.\tag{1.7.1}$$

直接由范数公理$(\mathrm{N}_1)\sim(\mathrm{N}_3)$及定义式(1.7.1)推出,距离有性质:

(M_1)**对称性**: $d(x,y)=d(y,x)$;

(M_2)**三角不等式**: $d(x,y)\leqslant d(x,z)+d(y,z)$;

(M_3)**正定性**: $d(x,y)\geqslant 0$; $d(x,y)=0\Leftrightarrow x=y$.

如果撇开距离的定义式(1.7.1),仅仅以性质(M_1)~(M_3)作为出发点,就达到公理化的距离定义.

定义 1.7.1 设 X 是任一非空集.若对 X 中任一对元 x,y,指定了一个实数 $d(x,y)$,称为 x 与 y 之间的距离,它满足**距离公理**(M_1)~(M_3),则称 $d(\cdot,\cdot)$ 为 X 上的一个**度量**或**距离**,称 X 为**度量空间**;当要明确指出度量 d 时,将空间 X 写作 (X,d).

由此定义,任何赋范空间依式(1.7.1)所定义的度量 d 是一个度量空间,称 d 为由范数导出的度量.当将一赋范空间看作度量空间而未另作说明时,总认定其中使用由范数导出的度量.实际上,赋范空间的任何非空子集依距离(1.7.1)式亦是度量空间.应用上常用的度量空间可能是某个赋范空间的子集,但亦可能通过其他途径引出.关键的一点是,度量空间不必具有任何代数结构,因而比赋范空间具有更大的一般性.但因此却远不能达到如赋范空间中那样丰富的结果.

设 (X,d) 是给定的度量空间.我们感兴趣的是,赋范空间中通过距离(1.7.1)式所定义的哪些概念及相应的结论可推广于度量空间.要回答这一问题,就得对§1.1~§1.4 中的内容作一番仔细审察,从其中挑出那些可推广到度量空间的结论.这件事做起来并不十分困难,但也难以从中感受到开辟新领域的乐趣.我们不打算去描述这一全过程,只是给出最初的几步.

若 $\{x_n\}\subset X$ 是一序列,$x\in X$,$d(x_n,x)\to 0(n\to\infty)$,则说 $\{x_n\}$ **度量收敛**于 x,记作 $x_n\to x(n\to\infty)$.若

$$\lim_{m,n} d(x_m,x_n)=0 \tag{1.7.2}$$

(对照式(1.1.9)),则称 $\{x_n\}$ 为 Cauchy 列.若 X 中所有的 Cauchy 列收敛,则称 X 为**完备度量空间**.Banach 空间中的任何非空闭子集都是完备度量空间.对应于式(1.3.1),式(1.3.2),度量空间 X 中的开球与闭球分别定义为

$$B_r(a)=\{x\in X:d(a,x)<r\}$$

与

$$\overline{B}_r(a)=\{x\in X:d(a,x)\leqslant r\}.$$

由此进而将§1.3 与§1.4 中的种种概念过渡到度量空间,几乎是例行的推广.要作的往往是将 $\|x-y\|$ 换成 $d(x,y)$,将"Banach 空间的闭子集"换成完备度量空间而已.不过,与其说说而已,不如实际动手完成某些推广.这件事你不妨试着自

己来做．作为示范，我们推广定理 1.4.4 如下．

定理 1.7.2　设 X,Y 是度量空间，其中 X 是紧的，$F:X\to Y$ 与 $f:X\to\mathbf{R}$ 连续，则有以下结论：

(i) FX 是 Y 中的紧集，因而有界；F 在 X 上一致连续，即 $\forall\varepsilon>0,\exists\delta>0$，当 $x,y\in X,d(x,y)<\delta$ 时 $d(Fx,Fy)<\varepsilon$；

(ii) f 在 X 上取得最大值与最小值．

定理中涉及度量空间中的紧集、有界集与连续性概念，参照 §1.3 与 §1.4，你应能自行给出定义．

证　(i) 对于 FX 为紧集的证明，与定理 1.4.4 中证 FD 紧无任何不同．若 F 在 X 上非一致连续，则对某个 $\varepsilon>0$，存在 $x_n,y_n\in X(n=1,2,\cdots)$，使得 $d(x_n,y_n)\to 0$，而 $d(Fx_n,Fy_n)\geqslant\varepsilon(\forall n\in\mathbf{N})$．因 X 是紧的，不妨设 $x_n\to x\in X$，因而亦有 $y_n\to x(n\to\infty)$，但这推出

$$\varepsilon\leqslant\lim_n d(Fx_n,Fy_n)=d(Fx,Fx)=0,$$

得出矛盾．因此 F 必定一致连续．

为证(ii)，几乎可以照抄定理 1.4.4(ii)的相应证明，因此可以免去．　□

当然，那些实质上依赖于一定代数关系的概念与结论，对于度量空间就不再有意义．例如，"基本集"的概念就不适用于度量空间．

无需具有代数结构，固然使得度量空间有更具普遍性的优点，但亦使它失去常识中熟悉的空间形象．例如，整数集 $\mathbf{Z}$ 依通常度量是一个度量空间，但在直观上，我们很难将它与 $\mathbf{R}^n$ 这样的"更地道"的空间等量齐观．在度量空间 $\mathbf{Z}$ 中，球

$$\bar{B}_1(0),\quad B_{3/2}(0),\quad B_2(0)$$

竟无区别，且 $\bar{B}_1(0)\neq\overline{B_1(0)}$．从习惯于赋范空间的眼光看来，这些事实几乎是不可思议的．因此，在处理度量空间问题时，应与来自平常空间的几何直观保持适当的距离，以免陷入谬误．无疑，缺少代数结构是抽象度量空间的一个缺点．因此，应用上最重要的度量空间，还是那些同时是向量空间且其线性运算依度量连续的度量空间．除了赋范空间之外，这类空间的最重要例子就要算下面即将介绍的赋准范空间了．

§1.7.2　赋准范空间

我们将定义 1.1.1 修改如下．

定义 1.7.3　设 X 是 $\mathbf{K}$ 上的向量空间．若对每个 $x\in X$，指定了一个实数 $\|x\|$，称为 x 的**准范数**，它满足范数公理(N_2)，(N_3)(依定义 1.1.1)及修正的(N_1)：

$$(\mathrm{N}'_1)\ \| -x \| = \| x \|;\ \lim_{\alpha_n \to 0} \| \alpha_n x \| = 0 = \lim_{\| x_n \| \to 0} \| \alpha x_n \|$$
$$(x, x_n \in X, \alpha_n \in \mathbf{K}),$$

则称 X 为 $\mathbf{K}$ 上的**赋准范空间**.

若 X 是赋准范空间,仍然依公式(1.7.1)定义距离 d, 则显然 (X,d) 是度量空间,称 d 为由准范数 $\|\cdot\|$ 导出的度量;当 (X,d) 完备时称 X 为 **Fréchet 空间**. 当将赋准范空间(以及其中任何非空子集)看作度量空间时,通常总使用由准范数导出的度量. 赋准范空间 X 中的度量收敛,就是依准范数收敛(对照定义 1.1.3):

$$d(x_n, x) \to 0 \Leftrightarrow \| x_n - x \| \to 0 \quad (n \to \infty).$$

如同在赋范空间中一样,如上的收敛有性质:

$$x_n \to x, \quad y_n \to y \Rightarrow x_n + y_n \to x + y; \tag{1.7.3}$$

$$\alpha_n \to \alpha, \quad x_n \to x \Rightarrow \alpha_n x_n \to \alpha x. \tag{1.7.4}$$

式(1.7.3)是很简单的;证明式(1.7.4)却不像赋范空间中那么容易(参看 §1.1.2 与文献[4]第 313 页). 一般说来,将范数公理(N_1)替换成(N'_1)之后,赋范空间中很多论证因而随之失效.

显然赋范空间必为赋准范空间,因而 Banach 空间也是 Fréchet 空间. 我们更关心的是那些不可赋范的 Fréchet 空间,下面就是一个典型例子.

例 1.7.4 $\mathbf{R}$ 上的连续函数之全体 $C(\mathbf{R})$ 显然是一向量空间. 对任给 $u \in C(\mathbf{R})$, 定义

$$\begin{cases} \| u \|_n = \sup\limits_{|x| \leqslant n} | u(x) | \quad (n \in \mathbf{N}), \\ \| u \| = \sum\limits_{n=1}^{\infty} \dfrac{1}{2^n} \dfrac{\| u \|_n}{1 + \| u \|_n}. \end{cases} \tag{1.7.5}$$

今验明 $\|\cdot\|$ 是 $C(\mathbf{R})$ 上的一个准范数. 范数公理$(\mathrm{N}_2)(\mathrm{N}_3)$的验证是容易的,三角不等式基于不等式

$$\frac{|u+v|}{1+|u+v|} \leqslant \frac{|u|}{1+|u|} + \frac{|v|}{1+|v|}.$$

其次显然 $\| -u \| = \| u \|$. 为验证公理(N'_1),首先指出:若 $\{u_k\} \subset C(\mathbf{R})$, 则

$$\begin{aligned} \| u_k \| \to 0 &\Leftrightarrow \| u_k \|_n \to 0 \qquad (k \to \infty, \forall\, n \in \mathbf{N}) \\ &\Leftrightarrow \text{在} [-n, n] \text{ 上 } u_k(x) \rightrightarrows 0 \quad (k \to \infty, \forall\, n \in \mathbf{N}). \end{aligned} \tag{1.7.6}$$

由此容易推出：若 $u, u_k \in C(\mathbf{R}), \alpha_k \in \mathbf{K}(k = 1,2,\cdots)$，则 $\alpha_k \to 0 \Rightarrow \| \alpha_k u \| \to 0$；$\| u_k \| \to 0 \Rightarrow \| \alpha u_k \| \to 0 (k \to \infty)$. 因此 $\| \cdot \|$ 是一个准范数；式(1.7.6)表明 $C(\mathbf{R})$ 中的度量收敛就是有限区间上的一致收敛，简称为**紧一致收敛**. 事实上，式(1.7.6)显然推出：

$$\| u_k \| \to 0 \Leftrightarrow \text{在任何紧集 } K(\subset \mathbf{R}) \text{ 上 } u_k \rightrightarrows 0 \quad (k \to \infty).$$

若 $\{u_k\} \subset C(\mathbf{R})$ 是一 Cauchy 列，即

$$\lim_{k,l} \| u_k - u_l \| = 0,$$

则

$$\lim_{k,l} \| u_k - u_l \|_n = 0 \quad (\forall n \in \mathbf{N}),$$

这表明限制在区间 $[-n, n]$ 上，$\{u_k\}$ 依 sup 范数是 Cauchy 列. 于是存在 $u \in C(\mathbf{R})$，使得在每个区间 $[-n, n] (n \geqslant 1)$ 上 $u_k \rightrightarrows u$，即 $\| u_k - u \| \to 0 (k \to \infty)$. 这表明 $C(\mathbf{R})$ 是完备的，从而是一 Fréchet 空间.

构成 Fréchet 空间的以上方法可自然地推广到 $C^m(\Omega)$，此处 $\Omega \subset \mathbf{R}^n$ 是任一开集(包括 $\Omega = \mathbf{R}^n$)，$0 \leqslant m \leqslant \infty$，使得在 $C^m(\Omega)$ 中序列 $\{u_k\}$ 度量收敛于零意味着：当 $|\alpha| \leqslant m$ (当 $m = \infty$ 时是对任给 $\alpha \in \mathbf{Z}_+^n$)时，$\partial^\alpha u_k$ 在 Ω 内紧一致收敛于零，即对任一紧集 $K \subset \Omega$，在 K 上 $\partial^\alpha u_k \rightrightarrows 0 (k \to \infty)$. 对于这样的空间，准范数的具体表达式可能有些复杂，但许多结论仅依赖于 Fréchet 空间的一般性质，而与准范数的具体表达式无关.

§1.7.3 拓扑空间

尽管度量空间已具有很大的一般性，但它仍然不足为用. 即使在看来很初等的问题上，度量空间的局限性亦能表现出来. 例如，以 F 记 $\mathbf{R}$ 上的实函数之全体，任给 $u, u_n \in F(n = 1,2,\cdots)$，定义

$$u_n \to u \Leftrightarrow u_n(x) \to u(x) \quad (n \to \infty, x \in \mathbf{R}),$$

即 F 中的收敛就是处处收敛或**点态收敛**. 这种收敛性当然极为简单初等，且不乏其用，但它却不能由任何度量定义①.

这样，我们又走到了一个进入新的空间领域的路口，路标指向拓扑空间②，它

① 这一结论的严格证明依赖于较深入的知识.

② 实际上，介于度量空间与拓扑空间之间，还有一些中间类型的抽象空间，但其重要性不及拓扑空间与度量空间，不必太予重视.

比度量空间还要宽泛得多，足以满足多种涉及收敛性的数学问题之需要.

回忆一下§1.3.1中关于点集的讨论.若以 τ 记一赋范空间 X 中的开集之全体，则它有以下性质(参看命题1.3.4)：

(O_1) $X,\varnothing\in\tau$；

(O_2) $\{A_i: i\in I\}\subset\tau\Rightarrow\bigcup A_i\in\tau$；

(O_3) $\{A_i: 1\leqslant i\leqslant n\}\subset\tau\Rightarrow\bigcap A_i\in\tau$.

上述结论同样也适用于度量空间.我们自然推想，这些性质或许根本不必与范数或度量发生必然联系，而应成为一个更一般的概念的出发点.这就导致以下定义.

定义1.7.5 设 X 是任一非空集.若已指定了一个集族 $\tau\subset 2^X$①，每个 $A\in\tau$ 称为开集，τ 满足**开集公理**$(O_1)\sim(O_3)$，则称 τ 为 X 上的一个**拓扑**，而称 X 为**拓扑空间**；当要明确指出拓扑 τ 时写作 (X,τ).

据此定义，任何赋范空间都是拓扑空间，其中的拓扑(即开集族)称为**范数拓扑**.同样地，任何度量空间都是拓扑空间，其中的拓扑称为**度量拓扑**.因此，赋范空间的任何非空子集是拓扑空间.特别，$\mathbf{R}^n$ 的任何非空子集(其度量由Euclid范数导出)是拓扑空间，其中的拓扑称为**通常拓扑**(这自然意味着在 $\mathbf{R}^n$ 中还可以考虑其他拓扑).

拓扑空间远比度量空间(更不必说赋范空间)更一般.要举出非度量空间的拓扑空间例子简直易如反掌.例如，设 X 是任何多于一点的集，令 $\tau=\{X,\varnothing\}$，则 τ 平凡地满足开集公理$(O_1)\sim(O_3)$，这样的 τ 就称为**平凡拓扑**.平凡拓扑绝不可能是度量拓扑.事实上，若拓扑 τ 由度量 d 导出，$x,y\in X$，$x\neq y$，则 $r\triangleq d(x,y)>0$，$B_{r/2}(x)$ 与 $B_{r/2}(y)$ 都应属于 τ，而显然 $B_{r/2}(x)\cap B_{r/2}(y)=\varnothing$，故不可能有 $\tau=\{X,\varnothing\}$.

一个颇具吸引力的问题是，能否在拓扑空间中展开你在§1.1～§1.4中所看到的那些内容？考虑到除抽象的公理$(O_1)\sim(O_3)$之外，现在别无其他结构可用，拓扑空间理论的展开似乎充满困难且内容贫乏.然而，实际上开集公理$(O_1)\sim(O_3)$蕴涵了一个异常丰富的理论，而且其基本特征非常接近于赋范空间中的点集论，以至无需太多的改动，即可将§1.3～§1.4中的许多内容移植到拓扑空间.当然这一移植过程仍然有不少困难需要克服，而且深入的展开将遇到更多的困难.所有这些都只能从拓扑学的专书中获得充分了解.本节充其量走出开头的几步，以使读者获得最初步的印象.

以下设 (X,τ) 是一取定的拓扑空间.

首先考虑收敛性.在一赋范空间中，$x_n\to x(n\to\infty)$ 可刻画为

---

① 2^X 表示 X 的子集之全体，称为 X 的**幂集**.注意 $X,\varnothing\in 2^X$，因而 2^X 必定非空.

$$\forall \varepsilon > 0, \exists N > 0, \forall n \geqslant N, \quad 有\ x_n \in B_\varepsilon(x).$$

在拓扑空间 X 中无球可用,我们用开集来代替球:在 X 中 $x_n \to x(n \to \infty)$ 界定为

$$任给包含\ x\ 的开集\ V, \exists N > 0, \forall n \geqslant N, 有\ x_n \in V. \qquad (1.7.7)$$

为行文方便,下面以 $\mathcal{N}_x$ 记含 x 的开集之全体,其中的集称为 x 的开邻域.

仿照极限的定义,以开集代替球,可将点集论的许多基本概念推广于拓扑空间(参看定义1.3.1,1.3.3,1.3.7,1.4.10).

定义 1.7.6 设 $A \subset X, x \in X$. 若存在 $V \in \mathcal{N}_x$,使 $V \subset A$,则称 x 为 A 的内点;若对任给 $V \in \mathcal{N}_x$,有 $A \cap V \neq \varnothing$,则称 x 为 A 的触点.分别以 A° 与 $\bar{A}$ 记 A 的内点与触点之全体,二者分别称为 A 的内部与闭包.若 $A = \bar{A}$,则称 A 为闭集;若 $\bar{A} = X$,则称 A 为稠集,若 $(\bar{A})^\circ = \varnothing$,则称 A 为疏集.

边界、导集、可分集、第一纲集等概念的定义如同赋范空间中一样,不必一一写出.其次,关系式(1.3.3)、命题 1.3.2(ii)、命题 1.3.4 等均可推广于拓扑空间.至于"A 是开集$\Leftrightarrow A = A^\circ$",则是一个可证明的结论,而不像在赋范空间中一样是用来定义开集的条件.你应记住,拓扑空间中的开集是依公理给定而作为整个理论出发点的.另一个初看起来出人意外但实属本质的差别是,在拓扑空间中一般不能证明"$x \in \bar{A} \Leftrightarrow$ 存在序列 $\{x_n\} \subset A$ 使 $x_n \to x$".为了用极限来刻画闭包,必须推广序列概念,考虑某种称为**网**的广义序列的极限,此处已不宜深论了.顺便指出,正是诸如此类的"病态"现象,造成了拓扑空间理论深入展开的困难.

观察一下§1.3.2 中刻画连续性的式(1.3.5),就很容易理解以下定义.

定义 1.7.7 设 X, Y 是两个拓扑空间,$F: X \to Y, x \in X, y = Fx$. 若对任给 $V \in \mathcal{N}_y$,存在 $U \in \mathcal{N}_x$,使得 $FU \subset V$,则说 F 在点 x 连续.若 F 在 X 上每点连续,则称 F 为连续映射.以 $C(X, Y)$ 记从 X 到 Y 的连续映射之全体.若 F 是一连续双射且 $F^{-1} \in C(Y, X)$,则称 F 为**同胚**.若存在从 X 到 Y 的同胚,则说拓扑空间 X 与 Y 互相同胚.

关于连续映射的定理 1.3.6 亦可用于拓扑空间.

因此处所定义的连续性涵盖了§1.3 中用范数所定义的连续性(参看定义1.3.5),故互相拓扑同构的赋范空间必定互相同胚.不过,同胚是比拓扑同构宽泛得多的概念.两个互相同胚的拓扑空间,初看起来其差别可能大得令人难以置信.例如,图 1-9 所示的 X 与 Y 作为 $\mathbf{R}^2$ 的两个拓扑子空间是互相同胚的,尽管二者似乎相差很远.这一类的事实促使人们去揭示隐藏在互相同胚的拓扑空间之后的共同属性,而这往往导向拓扑学的深刻结论.

对于紧性的刻画亦是拓扑空间理论的基本课题.但定理 1.4.3 已不能简单地推广于拓扑空间;在赋范空间中互相等价的条件,过渡到拓扑空间之后导向一些互有差异的拓扑概念.

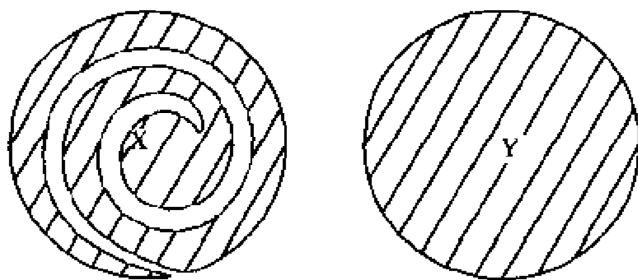


图 1-9

定义 1.7.8 设 X 是一拓扑空间，$A\subset X$. 若 A 在 X 中的任何开覆盖有有限子覆盖，则称 A 为紧集；当 $\bar{A}$ 是紧集时称 A 为相对紧集. 若 A 中任何序列有收敛子列且其极限属于 A，则称 A 为**序列紧集**. 若 X 本身是紧集(或序列紧集)，则称 X 为**紧空间**(或**序列紧空间**).

紧集未必是序列紧集，其逆也未必成立. 这种差异常常是处理拓扑空间问题的主要困扰之一.

现在考虑定理 1.4.4 的推广. 拓扑空间之间的映射已不再有一致连续概念，因此定理 1.4.4(i)不完全适用于拓扑空间. 不过，我们还是有以下结论.

定理 1.7.9 设 X, Y 是拓扑空间，其中 X 是紧的；$F: X\to Y$ 与 $f: X\to \mathbf{R}$ 连续，则有以下结论：

(i) FX 是 Y 中的紧集.

(ii) f 在 X 上取得最大与最小值.

证 (i) 设 $\{V_i: i\in I\}$ 是 FX 在 Y 中的开覆盖，则 $\{F^{-1}V_i\}$ 是 X 的开覆盖. 因 X 是紧的，故 $\{F^{-1}V_i\}$ 有有限子覆盖 $\{F^{-1}V_{i_k}\}$，因而 $\{V_{i_k}\}$ 是 $\{V_i\}$ 对 FX 的有限子覆盖. 这表明 FX 是紧集.

(ii) 令 $A=f(X)$，则 A 是 $\mathbf{R}$ 中的紧集，即有界闭集(参考定理 1.4.2). 令 $\alpha=\inf A$，$\beta=\sup A$. 因 A 是闭集，必定 $\alpha,\beta\in A$，故 α,β 分别为 f 在 X 上的最小值与最大值. □

进一步的深入将更加远离我们对于空间的直觉观念，在缺乏具体材料支撑的情况下，大概难以激起更大的兴趣.

不过，还有一个遗留问题有待回答：本段开头提到的空间 F 中的点态收敛能从某个拓扑导出吗？回答是肯定的，但这个拓扑的构造看起来还不简单：我们得实际构成一个集族 $\tau\subset 2^F$，验证它的确满足公理$(O_1)\sim(O_3)$；然后说明，F 中序列的点态收敛恰好是条件式(1.7.7)所刻画的那种收敛. 首先，对任给 $u\in F, x_1, \cdots, x_n\in\mathbf{R}, \varepsilon>0$，令

$$U(u,x_1,\cdots,x_n,\varepsilon)=\{v\in F:|u(x_i)-v(x_i)|<\varepsilon(1\leqslant i\leqslant n)\}. \tag{1.7.8}$$

定义 τ 为形如式(1.7.8)的集的并之全体，τ 就是我们所需要的拓扑．τ 应满足的各项条件逐一验证如下．

(i) 验证(O_1)．显然 F 是所有形如式(1.7.8)的集的并，故 $F\in\tau$．其次，"零个集的并"理解为空集，因此 $\varnothing\in\tau$．

(ii) (O_2)显然满足．

(iii) 验证(O_3)．只需验证：任意两个形如式(1.7.8)的集的交属于 τ．设 $A=U(u,x_1,\cdots,x_n,\varepsilon)$，$B=U(v,x_{n+1},\cdots,x_m,\delta)$，$u,v\in F$，$x_1,x_2,\cdots,x_m\in\mathbf{R}(n<m)$，$\varepsilon,\delta>0$．要证 $A\cap B\in\tau$，只要指明：$\forall w\in A\cap B$，存在一个形如式(1.7.8)的集 C，使得 $w\in C\subset A\cap B$．取定 $w\in A\cap B$．令

$$\eta=\min\{\varepsilon-|u(x_i)-w(x_i)|,\delta-|v(x_j)-w(x_j)|:1\leqslant i\leqslant n<j\leqslant m\};$$

$$C=U(w,x_1,x_2,\cdots,x_m,\eta),$$

只要证 $C\subset A\cap B$．任取 $z\in C$，有

$$\begin{aligned}|z(x_i)-u(x_i)|&\leqslant|z(x_i)-w(x_i)|+|w(x_i)-u(x_i)|\\&<\eta+|w(x_i)-u(x_i)|\leqslant\varepsilon\quad(1\leqslant i\leqslant n),\end{aligned}$$

这得出 $z\in A$；同理有 $z\in B$，故 $z\in A\cap B$．$C\subset A\cap B$ 得证．

这样，(F,τ) 是一拓扑空间，τ 是其中的开集族．

(iv) 验证拓扑空间 (F,τ) 中的序列收敛就是点态收敛．设 $u,u_n\in F(n=1,2,\cdots)$．若依条件式(1.7.7)有 $u_n\to u$，则 $\forall x\in\mathbf{R}$，$\forall\varepsilon>0$，因 $U\triangleq U(u,x,\varepsilon)$ 是一个含 u 的开集，故 $\exists N>0$，$\forall n\geqslant N$，有 $u_n\in U$，这正意味着

$$|u_n(x)-u(x)|<\varepsilon\quad(\forall n\geqslant N).$$

可见 $u_n(x)\to u(x)(n\to\infty,x\in\mathbf{R})$．反之，设 $u_n(x)$ 点态收敛于 $u(x)$．任给含 u 的开集 V，必有形如式(1.7.8)的集 $U=U(u,x_1,\cdots,x_n,\varepsilon)$，使 $U\subset V$．取 $N>0$，使得 $n\geqslant N$ 时

$$|u_n(x_i)-u(x_i)|<\varepsilon\quad(1\leqslant i\leqslant n),$$

则 $u_n\in U\subset V(\forall n\geqslant N)$，因此依条件式(1.7.7)有 $u_n\to u$．

习 题

1. 证明 $|d(x,A)-d(y,A)|\leqslant\|x-y\|$．

2. 证明命题 1.1.5.

3. 验证由式(1.1.11)所定义的范数满足范数公理.

4. 设 X_1, X_2 是两个 Banach 空间,证明 $X_1 \times X_2$ 依积范数是 Banach 空间.

5. 设 X, Y 是赋范空间, $T: X \to Y$ 是一个线性同构,证明 T 为拓扑同构的充要条件是:对任给 $\{x_n\} \subset X$, 有 $x_n \to x \Leftrightarrow Tx_n \to Tx$.

6. 证明任何多项式 $u(x) = \sum_0^n a_i x^i$ 满足不等式 $\sup\limits_{0 \leqslant x \leqslant 1} |u(x)| \leqslant C_n \sum |a_i|$, 其中常数 C_n 与 u 无关.

7. 设在 $[0,1]$ 上 $u_n \rightrightarrows u(n \to \infty)$, u_n 是次数不超过 5 的多项式. 证明 u 是次数不超过 5 的多项式.

8. 设 $\{x_1, \cdots, x_n\}$ 是赋范空间 X 中的线性无关子集. 证明对任何 $\lambda = (\lambda_1, \cdots, \lambda_n)^{\mathrm{T}} \in \mathbf{K}^n$ 有 $\alpha \sum |\lambda_i| \leqslant \| \sum \lambda_i x_i \| \leqslant \beta \sum |\lambda_i|$, 正常数 α, β 与 λ 无关.

9. 对任给 $x = (x_1, \cdots, x_n)^{\mathrm{T}} \in \mathbf{K}^n$, 令 $\|x\|_p = \left(\sum |x_i|^p\right)^{1/p} (1 \leqslant p < \infty)$, 说明 $(\mathbf{K}^n, \|\cdot\|_p)$ 是一个 Banach 空间. 其中的范数收敛是什么?

10. 设 $C_0(\mathbf{R}) = \{u \in C(\mathbf{R}) : \lim\limits_{|x| \to \infty} u(x) = 0\}$, 验证 $C_0(\mathbf{R})$ 依 sup 范数是 Banach 空间.

11. $[0,1]$上的多项式之全体依 sup 范数是否为赋范空间? 是否为 Banach 空间?

12. 设 $u, u_n \in L^p(J)(n = 1,2,\cdots)$, $1 \leqslant p \leqslant \infty$, $u_n \rightrightarrows u$, 证明 $u_n \xrightarrow{L^p} u$.

13. 设 $u(x) = \sum_0^n a_i x^i$, 理论结果断定 $u(x)$ 在$[0,1]$上无零点,但系数 a_i 由实验测定,可能有微小误差. 能否断定 $u(x)$ 在$[0,1]$上无零点? 为什么?

14. 设 A 是赋范空间 X 的真子空间 (即 $A \neq X$), 证明 $A^\circ = \varnothing$.

15. 设由实验确定,变量 u 与 x 的相关关系是一 2 次曲线 $u = u(x)(0 \leqslant x \leqslant 1)$. 如果实验条件有微小变动,能否保证结果依然为 2 次曲线? 为什么?

16. 设 A 是赋范空间 X 的子集,证明 $\{x : d(x, A) < r\}$ 是开集,画出该集的示意图.

17. 证明:在 $\mathbf{R}^2$ 中,有界$\Leftrightarrow$全有界.

18. 设 A, B 为紧集,证明存在 $a \in A, b \in B$, 使 $\|a - b\| = d(A, B)$.

19. 设 A 是紧集, B 是闭集, $A \cap B = \varnothing$, 证明存在 $r > 0$, 使得

$$\{x : d(x, A) < r\} \cap \{x : d(x, B) < r\} = \varnothing.$$

说明其直观意义.

20. 设 $\dim X = \infty$, A 是 X 的有限维子空间,证明存在 $x \in X$, 使得 $\|x\| = 1, d(x, A) = 1$.

21. 设 $J = [0,1]$, S 是 $C(J)$ 的单位球面. 证明 $f(u) = \int_0^1 |u(x)| \mathrm{d}x$ 在 $C(J)$ 上连续. $f(u)$ 在 S 上是否取得最大最小值?

22. 设 A 是 $C(J)$ 的有界子集, $J = [0,1]$, $\tilde{u}(x) = \int_0^x u(t) \mathrm{d}t$, 证明 $\{\tilde{u} : u \in A\}$ 在 $C(J)$ 中相对紧.

23. 设赋范空间 X 中存在线性无关无限集 $\{x_n : n \in \mathbf{N}\}$, 且 $X = \operatorname{span}\{x_n\}$, 证明 X 必不完

备.

24. 证明 $C(J)$ 中的非负函数之集是第二纲集.

25. 证明 $L^p(J)(1 \leqslant p < \infty)$ 中的非负函数之集是疏集.

26. 证明:在 Hilbert 空间中, $\{x_n\}$ 收敛⇔ $\lim\limits_{m,n}\langle x_m, x_n\rangle$ 存在且有限.

27. 证明:在内积空间中, $x_1, x_2, \cdots, x_n$ 线性无关⇔其 Gram 矩阵可逆.

28. 说明 $C[0,1]$ 中的 sup 范数不能由内积定义.

29. 证明:在内积空间中, $x \perp y \Leftrightarrow \forall \alpha \in \mathbf{K}$, 有 $\|x + \alpha y\| = \|x - \alpha y\|$, 解释其直观意义.若 $\mathbf{K} = \mathbf{R}$, 结论可作何修改?

30. 证明: $x \perp y \Leftrightarrow \forall \alpha \in \mathbf{K}$ 有 $\|x + \alpha y\| \geqslant \|x\|$, 说明其直观意义.

31. 用 Schimidt 正交化方法求 Legendre 多项式 $L_n, n = 0,1,2$.

32. 用 Schimidt 正交化方法求 Hermite 多项式 $H_n, n = 0,1,2$.

33. 在$[-1,1]$上求 $u(x) = 1/(1 + x^2)$ 的最佳均方逼近 2 次多项式.

34. 在 $[0,\pi]$ 上求 $u(x) = \sin x$ 的最佳均方逼近 2 次多项式.

35. 在$[0,1]$上求 $u(x) = e^x$ 的最佳均方逼近,使其在区间 $(1,1/3)$, $(1/3,2/3)$ 与 $(2/3,1)$ 上分别为常数.

第二章　线性算子与线性泛函

Banach 空间上的线性算子理论，构成泛函分析的核心内容，也是泛函分析应用于各个领域的主要工具. 从逻辑上说，线性算子理论无疑是有限维向量空间中线性变换理论的自然扩展，因而不可避免地沿用线性代数中的某些思路与术语，以至在某种意义上可将线性算子理论看作**无限维空间上的线性代数学**. 在本章及下章中，我们将充分利用这一比拟所能带来的启示. 不过，应强调指出，无限维空间所固有的复杂性，使我们将面对一个艰深得多的理论.

本章中用到 X, Y, Z 时总假定为赋范空间，在同一场合用到同一数域 $\mathbf{K}$，大部分结果同时适用于 $\mathbf{K}=\mathbf{R}$ 与 $\mathbf{C}$ 两种情况. 一部分结果依赖于空间的完备性，只要碰到这种情况，总会明确指出并加以强调.

§2.1　有界线性算子

§2.1.1　线性算子

首先概述纯代数意义上的线性算子概念，这对于阅读本章是不可缺少的预备知识. 有关内容见于标准的线性代数教科书，此处只是最大限度地加以概括而已. 本段中只需假定 X, Y, Z 等是($\mathbf{K}$ 上的)向量空间就够了.

若一个映射 $T: X \to Y$ 满足

$$T(\alpha x + \beta y) = \alpha Tx + \beta Ty \quad (\alpha, \beta \in \mathbf{K}, x, y \in X), \tag{2.1.1}$$

则称 T 为从 X 到 Y 的**线性算子**①. 恒等式(2.1.1)意味着 T **保持线性运算**，它可推广为更一般的

$$T\Big(\sum_i \alpha_i x_i\Big) = \sum_i \alpha_i Tx_i, \tag{2.1.1$'$}$$

其中 $\sum \alpha_i x_i$ 是 X 中元的任一有限线性组合. 正是恒等式(2.1.1)(或(2.1.1)′)决定了线性算子具有一系列良好的代数性质，今将其最基本的性质汇总如下.

命题 2.1.1　设 $T: X \to Y$ 是一线性算子，则以下结论成立：

(i) 任给子空间 $A \subset X$ 与子空间 $B \subset Y$，TA 与 $T^{-1}B$ 分别为 Y 与 X 的子空间. 特别，$T(0) = 0$；$R(T) = TX$ 是 Y 的子空间(值域)；$N(T) \triangleq T^{-1}(0)$ 是 X

① 亦称为线性映射、线性变换、线性函数等；当 $Y = \mathbf{K}$ 时称为**线性泛函**.

的子空间(称为 T 的**核**或**零空间**).

(ii) 若向量组 $\{x_i\} \subset X$ 线性相关,则 $\{Tx_i\}$ 亦线性相关;若 A 是 X 的子空间且 $\dim A < \infty$, 则 $\dim TA \leqslant \dim A$.

(iii) T 是单射 $\Leftrightarrow N(T) = \{0\}$.

证明是很简单的(详见线性代数方面的一些书籍).

若 $Tx \equiv 0 \in Y (x \in X)$, 则称 T 为**零算子**,就记作 0;若 $Tx \equiv \alpha x (x \in X)$, $\alpha \in \mathbf{K}$ 为常数,则称 T 为**纯量算子**(或相似变换,若 $\alpha \neq 0$),记作 αI, 当 $\alpha = 0$ 与 1 时 αI 分别为零算子 0 与单位算子 I.

对线性算子可定义两种自然的运算:线性运算与乘法.若 $T, S: X \to Y$ 是线性算子, $\alpha, \beta \in \mathbf{K}$, 则 $\alpha T + \beta S: X \to Y$ 是一个线性算子,它定义为

$$(\alpha T + \beta S)x = \alpha Tx + \beta Sx \quad (x \in X). \tag{2.1.2}$$

若 $R: Y \to Z$ 是另一个线性算子,则由

$$(RT)x = R(Tx) \quad (x \in X) \tag{2.1.3}$$

定义出一个线性算子 $RT: X \to Z$, 称它为 R 与 T 的乘积.实际上,线性算子的乘积就是它们的复合.容易验证,如上定义的运算有以下性质:

$$\begin{cases} R(T + S) = RT + RS, \\ (R + R_1)T = RT + R_1 T; \end{cases} \quad (\text{分配律})$$

$$Q(RT) = (QR)T; \qquad (\text{结合律})$$

$$\alpha(RT) = (\alpha R)T = R(\alpha T), (\alpha \in \mathbf{K})$$

只要以上等式的一端有意义.若线性算子 $T: X \to Y$ 为双射,则称它为线性同构,此时其逆映射 $T^{-1}: Y \to X$ 亦为线性算子. T 是线性同构的充要条件是,存在线性算子 $S: Y \to X$, 使得

$$ST = I_X, \quad TS = I_Y. \tag{2.1.4}$$

§2.1.2　有界线性算子

对于线性算子 $T: X \to Y$, 经常需要估计 $\|Tx - Ty\|$, 而且希望能估计得尽可能准确.因 $Tx - Ty = T(x - y)$, 实际上只需考虑估计 $\|Tx\| (x \in X)$. 这种估计通过考虑比值 $\|Tx\| / \|x\| (x \neq 0)$ 来解决.例如,若能断定 $\|Tx\| / \|x\| \leqslant k (0 \neq x \in X)$, 则从 $\|x\| \leqslant 1$ 可得出估计 $\|Tx\| \leqslant k$. 以上考虑导向如下定义.

定义 2.1.2　设 $T: X \to Y$ 是一个线性算子.令

$$\|T\| = \sup_{x\neq 0}\|Tx\|/\|x\|. \tag{2.1.5}$$

若 $\|T\| < \infty$，则称 T 为从 X 到 Y 的**有界线性算子**，且称 $\|T\|$ 为 T 的**算子范数**，简称为范数. 若 $\|T\| = \infty$，则称 T 为**无界算子**.

约定以 $L(X,Y)$ 记从 X 到 Y 的有界线性算子之全体，$L(X,X)$ 就简写作 $L(X)$.

直接从定义易推出，线性算子 $T: X\to Y$ 的有界性可等价地刻画为：

(i) $\exists k>0, \forall x\in X$，有 $\|Tx\|\leqslant k\|x\|$；或

(ii) T 映 X 中的有界集为 Y 中的有界集."有界算子"一词，即由此而来.

若 $T\in L(X,Y)$，则直接由式(2.1.5)得出，对任给 $x\in X$ 有

$$\|Tx\|\leqslant\|T\|\,\|x\|. \tag{2.1.6}$$

这就是给出 $\|Tx\|$ 估计的一般不等式，其应用极为广泛，可以说是泛函分析中用得最频繁的不等式之一.

对于 $T\in L(X,Y)$，算子范数 $\|T\|$ 是一个最重要的数量指标. 它的直观意义是什么呢？为看得更清楚些，我们给出式(2.1.5)的几个等价形式：

$$\|T\| = \sup_{\|x\|=1}\|Tx\| \tag{2.1.7}$$

$$= \sup_{\|x\|\leqslant 1}\|Tx\| \tag{2.1.8}$$

$$= \inf\{k\geqslant 0: \|Tx\|\leqslant k\|x\|\ (\forall x\in X)\}. \tag{2.1.9}$$

为验证式(2.1.7)～式(2.1.9)，首先由式(2.1.5)，式(2.1.6)易推出

$$\|T\| = \sup_{\|x\|=1}\|Tx\|\leqslant\sup_{\|x\|\leqslant 1}\|Tx\|\leqslant\|T\|,$$

这得出等式(2.1.7)与(2.1.8). 若以 ρ 记等式(2.1.9)之右端，则由式(2.1.6)推出 $\rho\leqslant\|T\|$. 另一方面，若 $\|Tx\|\leqslant k\|x\|\ (\forall x\in X)$，则

$$\|T\| = \sup_{\|x\|=1}\|Tx\|\leqslant k,$$

这推出 $\|T\|\leqslant\rho$. 因此 $\|T\|=\rho$，即式(2.1.9)成立. 因取 $k=\|T\|$ 使式(2.1.9)中的下确界达到，故式(2.1.9)中的 inf 可改为 min.

现在利用算子范数公式(2.1.5)与式(2.1.7)～式(2.1.9)从不同角度阐明 $\|T\|$ 的直观意义. 因 $\|Tx\|/\|x\|$ 是 Tx 与 x 两者的"长度"之比，故式(2.1.5)表明 $\|T\|$ 是变换 T 的"最大伸张系数"①. 其次，X 中的闭单位球 $\overline{B}_1(0)$ 的象 $T\overline{B}_1(0)$ 包含于 Y 中某个以 0 为心的最小闭球，而式(2.1.8)表明此闭球的半径正

① 这只是一种启发性的说法，"最大"一词并不意味着式(2.1.5)中的上确界一定达到.

是 $\|T\|$；式(2.1.7)则表明 $TS_1(0)$ 已"触到"球面 $S_{\|T\|}(0)$，此处以 $S_r(0)$ 记球面 $\{x : \|x\| = r\}$. 最后，注意不等式 $\|Tx\| \leqslant k\|x\|(\forall x \in X)$ 给出对 $\|Tx\|$ 的估计；其中 k 愈小，所得的估计就愈精确. 于是式(2.1.9)表明，不等式(2.1.6)给出对 $\|Tx\|$ 的某种最佳估计. 对于有界线性算子的应用，关于 $\|T\|$ 的最后这个解释或许是最有意义的.

纯量算子 αI 显然有范数 $|\alpha|$；特别 $\|I\| = 1$. 下面考虑一个不那么平凡的解释性例子. 下节中将系统地考察更多的具体线性算子.

例 2.1.3　设 $J = [a,b](a < b)$，$\varphi \in C(J)$. 定义

$$Tu(x) = \varphi(x)u(x)^{①} \quad (x \in J, u \in C(J)),$$

T 显然是从 $C(J)$ 到自身的线性算子. 今求出 $\|T\|$，下面所用的方法有一定普遍性，值得你仔细体会. 依前面对算子范数的解释，我们要求得对 $\|Tu\|_0$ 的一个"最佳估计". 首先作一初步估计：

$$\|Tu\|_0 = \sup_{x\in J}|\varphi(x)u(x)| \leqslant \sup_{x\in J}|\varphi(x)| \sup_{x\in J}|u(x)|$$

$$= \|\varphi\|_0 \|u\|_0 \qquad (\forall u \in C(J)),$$

由此得出 $\|T\| \leqslant \|\varphi\|_0$（依式(2.1.9)). 为验证上述估计实际上是"最佳的"，选择特殊的 $u \in C(J)$ 进行检验，通常取满足 $\|u\|_0 = 1$ 的 u，具体选定依赖于直接观察(这往往是困难之所在). 此处倒很简单：取 $u(x) \equiv 1$，则 $\|u\|_0 = 1$，于是

$$\|\varphi\|_0 = \|Tu\|_0 \leqslant \|T\| \|u\|_0 = \|T\|,$$

因而 $\|T\| = \|\varphi\|_0$.

命题 2.1.4　设 $T : X \to Y$ 是一个线性算子，则 T 有界 $\Leftrightarrow$ T 连续.

证　若 T 有界，则当 X 中 $x_n \to x(n \to \infty)$ 时，

$$\|Tx_n - Tx\| \leqslant \|T\| \|x_n - x\| \to 0 \quad (n \to \infty),$$

可见 T 连续. 反之，若 T 无界，则必存在 $\{x_n\} \subset X$，使得 $\|x_n\| = 1$，$\|Tx_n\| > n(n = 1,2,\cdots)$. 这推出

$$\|n^{-1}x_n\| \to 0 \text{ 而 } 1 < \|T(n^{-1}x_n)\| \not\to 0,$$

因而 T 在 $x = 0$ 处不连续. □

---

① 左端表示函数 Tu，即算子 T 作用于 u 而得之函数. 因此，$Tu(x)$ 的更准确的写法应是 $(Tu)(x)$. 不过，只要不致误解，就常用较简单的写法 $Tu(x)$，此处及今后均如此.

注意，在命题 2.1.4 之证中实际上证明了：T 有界$\Leftrightarrow$ T 在 $x=0$ 处连续.

若 $T: X \to Y$ 是线性同构，则 §1.1 中条件(1.1.12)式等价于 T 与 T^{-1} 皆有界.这结合命题 2.1.4 推出：T 是拓扑同构$\Leftrightarrow$ T 与 T^{-1} 皆连续(即 T 为同胚，参看定义1.7.7).

命题 2.1.4 的意义在于，它说明有界线性算子不仅与线性运算可交换(这表为式(2.1.1)′)，而且亦与极限运算可交换.因此，若 $T \in L(X,Y)$，$\{x_n\} \subset X$，则

$$T\left(\sum_n x_n\right) = \sum_n Tx_n,$$

只要其中级数 $\sum x_n$ 收敛.今后将不加解释地运用类似的等式.

利用定理 1.1.7 不难证明，有限维赋范空间上的线性算子必定连续，因而是有界的.正是有限维空间上线性算子的良好性质给我们造成了一种印象，似乎线性算子理所当然地连续.实际上，在无限维空间中并非如此，不少常用且看来很“简单”的线性算子是不连续的，因而很难将不连续线性算子视为病态.甚至可以说，在无限维空间中，无界线性算子的出现几乎有某种普遍性，举出个别例子更是易如反掌.下面就是一个简单例子.

例 2.1.5 设 $J=[0,\pi]$，在 $C^1(J)$ 与 $C(J)$ 中都采用 sup 范数.显然

$$T = \frac{\mathrm{d}}{\mathrm{d}x}: C^1(J) \to C(J), \quad u \to u' \tag{2.1.10}$$

是一个线性算子.令 $u_n(x) = \sin nx$，则 $\| u_n \|_0 = 1$，而 $\| u' \|_0 = n$，可见 T 是无界算子.

从实际应用的需要来看，无界线性算子如同有界线性算子一样具有重要价值，因而不可当作例外排除(你能排除如式(2.1.10)这样的微分算子吗).不过，有界线性算子理论相对来说要简单些，因而自然成为优先考虑的对象.在带有入门性质的本书中，我们将主要考虑有界线性算子.

§2.1.3 有界线性算子的运算与扩张

我们要证明，“算子范数”事实上是一个范数.

命题 2.1.6 $L(X,Y)$ 依算子范数是一个赋范空间；当空间 Y 完备时，$L(X,Y)$ 为 Banach 空间.

证 要证的有以下三点：

(i) $L(X,Y)$ 依由式(2.1.2)定义的线性运算是 $\mathbf{K}$ 上的向量空间.这是显然的.

(ii) 算子范数满足范数公理(N_1)～(N_3)(依定义(1.1.1).(N_1)与(N_3)是明显的，今验证(N_2).任给 $T,S \in L(X,Y)$，$x \in X$，有

$$
\begin{aligned}
\|(T+S)x\| &\leqslant \|Tx\| + \|Sx\| \\
&\leqslant \|T\|\,\|x\| + \|S\|\,\|x\| \qquad \text{(用式(2.1.6))} \\
&= (\|T\| + \|S\|)\|x\|,
\end{aligned}
$$

这推出 $\|T+S\| \leqslant \|T\| + \|S\|$(依式(2.1.9)).

(iii) 设 Y 完备,证 $L(X,Y)$ 完备. 设 $\{T_n\} \subset L(X,Y)$ 是 Cauchy 列,即

$$
\lim_{m,n} \|T_m - T_n\| = 0. \tag{2.1.11}
$$

$\forall x \in X$, 用式(2.1.6)与式(2.1.11)推出

$$
\|T_m x - T_n x\| \leqslant \|T_m - T_n\|\,\|x\| \to 0 \quad (m,n \to \infty),
$$

可见 $\{T_n x\}$ 是 Y 中的 Cauchy 列. 因 Y 完备,故 $\{T_n x\}$ 收敛,令

$$
Tx = \lim_n T_n x \quad (x \in X).
$$

这就得到一个算子 $T: X \to Y$, 它显然是线性的. 由

$$
\begin{aligned}
\|T_n x - Tx\| &= \lim_m \|T_n x - T_m x\| \\
&\leqslant \overline{\lim_m} \|T_n - T_m\|\,\|x\|
\end{aligned}
$$

推出

$$
\|T_n - T\| \leqslant \overline{\lim_m} \|T_n - T_m\|,
$$

从而

$$
\overline{\lim_n} \|T_n - T\| \leqslant \overline{\lim_n}\,\overline{\lim_m} \|T_n - T_m\| = 0. \quad \text{(用(11))}
$$

这表明 $\|T_n - T\| \to 0 (n \to \infty)$, $T = T_n + (T - T_n) \in L(X,Y)$. $L(X,Y)$ 的完备性得证. □

在定义一个有界线性算子时,常用到以下定理.

定理 2.1.7(扩张定理) 设 D 是 X 的稠密子空间, $T \in L(D,Y)$, Y 完备,则 T 可保持范数惟一地扩张到 X 上.

证 任给 $x \in X$, 取序列 $\{x_n\} \subset D$, 使得 $x_n \to x (n \to \infty)$. 由

$$
\|Tx_m - Tx_n\| \leqslant \|T\|\,\|x_m - x_n\| \to 0 \quad (m,n \to \infty)
$$

知 $\{Tx_n\}$ 是 Y 中的 Cauchy 列. 由 Y 完备得出 $\{Tx_n\}$ 收敛, 于是可定义

$$
T_1 x = \lim_n Tx_n. \tag{2.1.12}
$$

若另有 $\{y_n\} \subset D$ 满足 $y_n \to x (n \to \infty)$, 则合并 $\{x_n\}$ 与 $\{y_n\}$ 为序列

$$\{z_n\} = \{x_1, y_1, x_2, y_2, \cdots\}$$

之后仍有 $z_n \to x$，从而 $\{Tz_n\}$ 收敛. 这就推出

$$\lim_n Tx_n = \lim_n Tz_n = \lim_n Ty_n,$$

可见式(2.1.12)与 $\{x_n\}$ 的选择无关，因而它确定地定义出一个算子 $T_1 : X \to Y$，T_1 显然是线性的. 若 $x \in D$，则在式(2.1.12)中取 $x_n \equiv x$ 得 $T_1 x = Tx$，可见 T_1 是 T 的扩张. 由式(2.1.12)推出

$$\| T_1 x \| = \lim_n \| Tx_n \| \leqslant \lim_n \| T \| \| x_n \| = \| T \| \| x \|,$$

可见 $\| T_1 \| \leqslant \| T \|$. 其次由式(2.1.8)显然有 $\| T \| \leqslant \| T_1 \|$，故 $\| T \| = \| T_1 \|$. 式(2.1.12)表明 T_1 由 T 惟一决定，故定理得证. □

定理 2.1.7 表明，为了定义一个 $T \in L(X, Y)$，只需在 X 的某个稠密子空间上定义 T，并验证其有界就行了. 这就简化了有界线性算子的构成. 实际上，还可以更进一步：只要在 X 的某个基本集 A 上定义 T，并验证

$$\left\| \sum_i \beta_i T a_i \right\| \leqslant k \left\| \sum_i \beta_i a_i \right\| \tag{2.1.13}$$

就行了，其中 $\sum \beta_i a_i$ 是 A 中元的任一有限线性组合，k 是与 $\sum \beta_i a_i$ 无关的正常数. 事实上，若

$$x = \sum_i \beta_i a_i = \sum_i \beta'_i a_i,$$

则由式(2.1.13)有

$$\left\| \sum_i \beta_i T a_i - \sum_i \beta'_i T a_i \right\| = \left\| \sum_i (\beta_i - \beta'_i) T a_i \right\|$$

$$\leqslant k \left\| \sum_i (\beta_i - \beta'_i) a_i \right\| = 0;$$

因此利用

$$T\left(\sum_i \beta_i a_i \right) = \sum_i \beta_i T a_i$$

将 $T : A \to Y$ 惟一地扩张为算子 $T : \mathrm{span} A \to Y$. 直接看出这样定义的 T 是线性的；不等式(2.1.13)表明 $\| T \| \leqslant k$. 于是应用定理 2.1.7 可将 T 扩张为 X 上的有界线性算子.

如上的扩张程序在泛函分析及其应用中被广泛采用. 试看一个简单而不乏启发性的例子. 任给 $\delta = (\alpha, \beta) \subset [a, b]$，定义

$$f(\chi_\delta) = \beta - \alpha.$$

若 $\delta_i = (\alpha_i, \beta_i) \subset [a, b]$ 是有限个区间，$\{\lambda_i\} \subset \mathbf{K}$，则

$$\left|\sum_i \lambda_i f(\chi_{\delta_i})\right| = \left|\sum_i \lambda_i (\beta_i - \alpha_i)\right| \leqslant (b - a)\left\|\sum_i \lambda_i \chi_{\delta_i}\right\|_0.$$

这样，f 就可扩张为 $B(J)(J = [a, b])$ 的某个子空间 X 上的有界线性泛函，X 以阶梯函数为其稠密子空间．显然 X 包含了 J 上的所有连续函数．你一定已经看出，$f(u)(u \in X)$ 就是 u 在 $[a, b]$ 上的积分：

$$f(u) = \int_a^b u(x)\mathrm{d}x.$$

这就得到定义积分的一种新方法，它并不直接用到任何“分割、求和、取极限”的通常程序．

若线性算子 $T: X \to Y$ 是单射(即 $N(T) = \{0\}$，参考命题 2.1.1)，则 $T^{-1}: \mathbf{R}(T) \to X$ 是一个确定的线性算子，当它有界时称为 T 的**有界逆**，并说 T 有有界逆．当需要判定一线性算子有有界逆时，经常用得着以下简单结果．

命题 2.1.8 线性算子 $T: X \to Y$ 有有界逆的充要条件是存在 $k > 0$，使得

$$\| Tx \| \geqslant k \| x \| \quad (x \in X). \tag{2.1.14}$$

证 若 $T^{-1} \in L(R(T), X)$，则 $\forall x \in X$，有

$$\| x \| = \| T^{-1}Tx \| \leqslant \| T^{-1} \| \| Tx \|.$$

必 $\| T^{-1} \| > 0$($X = \{0\}$ 这种情况不予考虑)，于是取 $k = \| T^{-1} \|^{-1}$ 使(2.1.14)满足．反之，若有 $k > 0$ 使式(2.1.14)满足，则显然 $Tx = 0 \Rightarrow x = 0$，因而 $N(T) = \{0\}$，$T^{-1}: R(T) \to X$ 有定义．$\forall x \in X$，由式(2.1.14)推出

$$\| T^{-1}(Tx) \| = \| x \| \leqslant k^{-1} \| Tx \|,$$

这表明 $T^{-1} \in L(R(T), X)$． □

当条件式(2.1.14)满足时，有些作者称 T 为**下有界算子**．注意这一概念与 T 本身是否有界无关．

§2.2 常用有界线性算子

本节系统地考虑若干常用的有界线性算子，它们一方面是解释概念的适当例子，同时也是应用泛函分析方法时经常使用的工具．

§2.2.1 矩阵

由线性代数熟知，矩阵是研究有限维空间中线性算子的有效工具．设 X, Y 是

有限维赋范空间，$\dim X = n, \dim Y = m, T \in L(X,Y)$. 分别取 X 的基 $\{e_j\}$ 与 Y 的基 $\{\epsilon_i\}$. 设

$$Te_j = \sum_i a_{ij}\epsilon_i \quad (1 \leqslant j \leqslant n),$$

则 T 完全由矩阵 $A = [a_{ij}] \in \mathbf{K}^{m\times n}$ 所确定. 若 $T, S \in L(X,Y)$ 分别对应矩阵 A, $B \in \mathbf{K}^{m\times n}$, $\alpha, \beta \in \mathbf{K}$, 则算子 $\alpha T + \beta S$ 恰对应矩阵 $\alpha A + \beta B$. 这样，线性算子空间 $L(X,Y)$ 线性同构于矩阵空间 $\mathbf{K}^{m\times n}$，因而对 $L(X,Y)$ 的研究可代之以对 $\mathbf{K}^{m\times n}$ 的研究. 这正是线性代数中熟知的观点.

任给 $A = [a_{ij}] \in \mathbf{K}^{m\times n}$，依矩阵乘法自然地定义一个线性算子：

$$\mathbf{K}^n \to \mathbf{K}^m, \quad x \to Ax, \tag{2.2.1}$$

其中 x 当作 $n\times 1$ 阶矩阵. 不妨用同一字母 A 表示算子(2.2.1)式，它也可以表成：

$$\begin{cases} y = Ax, \quad x = (x_j) \in \mathbf{K}^n, \quad y = (y_i) \in \mathbf{K}^m, \\ y_i = \sum_j a_{ij}x_j, \quad i = 1,2,\cdots,m. \end{cases} \tag{2.2.1$'$}$$

若在 $\mathbf{K}^n$ 中使用范数

$$\|x\|_p = \begin{cases} \left(\sum_j |x_j|^p\right)^{1/p}, & 1 \leqslant p < \infty, \\ \max_j |x_j|, & p = \infty, \end{cases} \tag{2.2.2}$$

则 $\mathbf{K}^n$ 可看作 l^p 的子空间，只要将 $x = (x_j) \in \mathbf{K}^n$ 等同于 l^p 中的元

$$(x_1, \cdots, x_n, 0, 0, \cdots)^{\mathrm{T}}.$$

通常称范数式(2.2.2)为 **p 范数**，采用 p 范数的 $\mathbf{K}^n$ 也记作 l_n^p. 相应地，算子 $A: l_n^p \to l_m^p$ (定义如(1))的范数记作 $\|A\|_p$，即(依式(2.1.8))

$$\|A\|_p = \sup_{\|x\|_p \leqslant 1} \|Ax\|_p. \tag{2.2.3}$$

$\|A\|_p$ 亦称为 A 的 p 范数. 以下结果在矩阵理论中是熟知的.

命题 2.2.1 设 $A = [a_{ij}] \in \mathbf{K}^{m\times n}$，则

$$\|A\|_1 = \max_j \sum_i |a_{ij}|; \tag{2.2.4}$$

$$\|A\|_\infty = \max_i \sum_j |a_{ij}| = \|A^{\mathrm{T}}\|_1; \tag{2.2.5}$$

$$\|A\|_2 = \max_j \sqrt{\lambda_j}, \quad \{\lambda_j\} \text{ 是 } A^{\mathrm{T}}A \text{ 的特征值之全体}. \tag{2.2.6}$$

对于无限维空间上线性算子的研究，矩阵不再是一个普遍有效的工具. 不过，我们仍可类比于有限维空间的情况，将某些无穷矩阵用于空间 l^p，得到类似于命

题 2.2.1 的结果.

以 $A = [a_{ij}](i,j = 1,2,\cdots)$ 记一个**无穷矩阵**,其中 $a_{ij} \in \mathbf{K}$. 仿照(1)′,形式地定义一个算子 $x \to Ax$:

$$\begin{cases} y = Ax, \quad x = (x_j), \quad y = (y_i), \\ y_i = \sum_j a_{ij}x_j, \quad i = 1,2,\cdots. \end{cases} \tag{2.2.7}$$

仍将式(2.2.7)所定义的算子记作 A,它在其有定义的集合上显然是线性的.不过,与有限维情况式(2.2.1)′不同的是,要使式(2.2.7)中的级数 $\sum_j a_{ij}x_j$ 收敛,且 y 属于适当的空间,对于 A 与 x 都需作一定限制.有趣的是,命题 2.2.1 的结论仍可(部分地)用于无穷矩阵.

命题 2.2.2 设算子 A 定义如式(2.2.7), $\|A\|_p$ 依式(2.2.3)(但假定其中 $x \in l^p$).

(i) 若 $\beta \triangleq \sup_j \sum_i |a_{ij}| < \infty$,则 $A \in L(l^1)$ 且 $\|A\|_1 = \beta$.

(ii) 若 $\beta \triangleq \sup_i \sum_j |a_{ij}| < \infty$,则 $A \in L(l^\infty)$ 且 $\|A\|_\infty = \beta$.

(iii) 若 $\beta \triangleq \left(\sum_{i,j} |a_{ij}|^2\right)^{1/2} < \infty$,则 $A \in L(l^2)$ 且 $\|A\|_2 \leqslant \beta$[①].

证 (i) 依处理例 2.1.3 的经验,我们必须证两个不等式: $\|A\|_1 \leqslant \beta$ 与 $\|A\|_1 \geqslant \beta$. 首先, $\forall x = (x_j) \in l^1$,有

$$\begin{aligned} \|Ax\|_1 &= \sum_i \left|\sum_j a_{ij}x_j\right| \leqslant \sum_i \sum_j |a_{ij}||x_j| \\ &= \sum_j |x_j| \sum_i |a_{ij}| \qquad \text{(交换求和顺序)②} \\ &\leqslant \beta \sum_j |x_j| = \beta \|x\|_1, \qquad \text{(依式(1.2.8))} \end{aligned}$$

这得出 $A \in L(l^1)$ 且 $\|A\|_1 \leqslant \beta$[③].

---

① 亦可能建立一个类似于式(2.2.6)的等式,但此处还缺少必要的知识来做到这一点.对于命题 2.2.3(iii)亦是如此.

② 由 Fubini 定理推出,当 $a_{ij} \geqslant 0$ 时恒有 $\sum_i \sum_j a_{ij} = \sum_j \sum_i a_{ij}$.

③ 更细致的论证应该首先说明,当 $x \in l^1$ 时 Ax 有定义,即 $\sum_j a_{ij}x_j$ 收敛.不过,推导 $\|Ax\|_1 \leqslant \beta\|x\|_1$ 的过程已显示出 $\sum_j |a_{ij}x_j| < \infty$. 因此,我们省略这一步.今后在类似情况下也采用从简的表述方式.

为证 $\|A\|_1 \geqslant \beta$，只需对任给 $\alpha < \beta$ 证 $\|A\|_1 > \alpha$（注意这是一种常用证法）. 取定一个 $\alpha < \beta$. 由 β 的定义，有下标 j_0，使 $\alpha < \sum_i a_{ij_0}$. 设 $\{e_j\}$ 是 l^1 的标准基（参考例 1.3.8(i)），则 $\|e_j\|_1 = 1$，Ae_j 是矩阵 A 的第 j 列. 于是

$$\alpha < \sum_i |a_{ij_0}| = \|Ae_{j_0}\|_1 \leqslant \|A\|_1 \|e_{j_0}\|_1 = \|A\|_1,$$

如所要证.

(ii) 大体上依照证(i)的思路：$\forall x = (x_j) \in l^\infty$，有

$$\begin{aligned}\|Ax\|_\infty = \sup_i \Big|\sum_j a_{ij}x_j\Big| &\leqslant \sup_i \sum_j |a_{ij}||x_j| \\ &\leqslant \sup_i \sum_j |a_{ij}| \|x\|_\infty = \beta \|x\|_\infty,\end{aligned}$$

这表明 $A \in L(l^\infty)$ 且 $\|A\|_\infty \leqslant \beta$. 其次，取定 $\alpha < \beta$，取 i_0 使 $\alpha < \sum_j |a_{i_0 j}|$. 令 $x_j = \operatorname{sgn} a_{i_0 j}$①，$x = (x_j)$，则 $\|x\|_\infty = 1$，

$$\begin{aligned}\alpha < \sum_j |a_{i_0 j}| &= \sum_j a_{i_0 j}x_j \\ &\leqslant \|Ax\|_\infty \leqslant \|A\|_\infty \|x\|_\infty = \|A\|_\infty.\end{aligned}$$

这推出 $\|A\|_\infty \geqslant \beta$，因此 $\|A\|_\infty = \beta$.

(iii) 任给 $x = (x_j) \in l^2$，有

$$\begin{aligned}\|Ax\|_2^2 &= \sum_i \Big|\sum_j a_{ij}x_j\Big|^2 \\ &\leqslant \sum_i \sum_j |a_{ij}|^2 \sum_j |x_j|^2 \quad (\text{用 §1.2(11)}) \\ &= (\beta\|x\|_2)^2,\end{aligned}$$

这推出 $\|Ax\|_2 \leqslant \beta\|x\|_2$. 因此 $A \in L(l^2)$ 且 $\|A\|_2 \leqslant \beta$. □

最简单的无穷矩阵莫过于对角形

$$D = \operatorname{diag}(a_1, a_2, \cdots, a_i, \cdots).$$

令 $a = (a_i)$. 则当 $a \in l^\infty$ 时由命题 2.2.2(i)(ii)有 $\|D\|_1 = \|D\|_\infty = \|a\|_\infty$；当 $a \in l^2$ 时由命题 2.2.2(iii)有 $\|D\|_2 \leqslant \|a\|_2$. 实际上，只要 $a \in$

① 严格地说，记号 sgn x 用于实数 x. 为方便起见，本书变通用之，作以下约定：若 $0 \neq x \in \mathbf{C}$，则令 $\operatorname{sgn} x = \bar{x}/|x|$；令 $\operatorname{sgn} 0 = 1$. 这样，恒有 $|\operatorname{sgn} x| = 1$，$x \operatorname{sgn} x = |x|$.

l^∞，就可直接推出 $\|D\|_2 = \|a\|_\infty$（参看本章习题 3）. 可见，命题 2.2.2(iii)中对 $\|A\|_2$ 的估计是比较粗疏的.

若 H 是以 $\{e_i : i \in \mathbf{N}\}$ 为标准正交基的 Hilbert 空间，$T \in L(H)$，$Te_j = \sum_i a_{ij} e_i$，则 T 完全由无穷矩阵 $A = [a_{ij}]$ 确定. 因此，对于可分 Hilbert 空间上有界线性算子的研究，矩阵仍不失为一个有用工具.

§2.2.2　积分算子

从由级数定义的算子(2.2.7)过渡到积分算子是极其自然的. 无穷矩阵 $A = [a_{ij}]$ 可以看作一个定义在 $\mathbf{N}\times\mathbf{N}$ 上的 2 元函数 $(i,j)\to a_{ij}$. 若以 $J=[a,b](a<b)$ 代替 $\mathbf{N}$，以 $J\times J$ 上的函数 $K(x,y)$ 代替 $A=[a_{ij}]$，以对 y 的积分代替对 j 求和，则得到对应于式(2.2.7)的如下积分算子：

$$Tu(x) = \int_a^b K(x,y)u(y)\mathrm{d}y \qquad (x \in J). \tag{2.2.8}$$

式(2.2.8)中的积分自然应理解为 Lebesgue 积分，函数 $K(x,y)$ 称为积分算子 T 的核或核函数. 类似于算子(2.2.7)式，为使式(2.2.8)中的积分存在（实际上只要求对几乎所有 $x\in J$ 积分存在），且函数 Tu 属于某个预定的函数空间，对于 $K(\cdot,\cdot)$ 与 u 均需有所限制. 与命题 2.2.2 对应的结果如下.

命题 2.2.3　设 $K(x,y)$ 是 $J\times J$ 上的可测函数，算子 T 依式(2.2.8)定义，约定 $\operatorname{ess\,sup}_x |\varphi(x)| = \|\varphi\|_\infty$（$L^\infty$ 范数又称为**本性上确界**）.

(i) 若 $\beta \triangleq \operatorname{ess\,sup}_{y\in J}\int_a^b |K(x,y)|\mathrm{d}x < \infty$，则 $T\in L(L^1(J))$ 且 $\|T\| = \beta$.

(ii) 若 $\beta \triangleq \operatorname{ess\,sup}_{x\in J}\int_a^b |K(x,y)|\mathrm{d}y < \infty$，则 $T\in L(L^\infty(J))$ 且 $\|T\| = \beta$.

(iii) 若 $\beta \triangleq \left(\int_a^b\int_a^b |K(x,y)|^2\mathrm{d}x\mathrm{d}y\right)^{1/2} < \infty$，则 $T\in L(L^2(J))$ 且 $\|T\| \leqslant \beta$.

如同在命题 2.2.2 中一样，在命题 2.2.3 中，对应于情况(i)～(iii)的 $\|T\|$ 可分别写作 $\|T\|_1$，$\|T\|_\infty$ 与 $\|T\|_2$.

作为例子，今将命题 2.2.3 用到简单的积分算子：

$$Tu(x) = \int_a^x u(y)\mathrm{d}y. \quad (x\in J) \tag{2.2.9}$$

只要取

$$K(x,y)=\begin{cases}1, & y\leqslant x,\\ 0, & y>x,\end{cases}$$

就可将式(2.2.9)写成标准形式(2.2.8)式.于是依命题2.2.3有:

$$\|T\|_1=\operatorname*{ess\,sup}_{y\in J}\int_y^b \mathrm{d}x=b-a;$$

$$\|T\|_\infty=\operatorname*{ess\,sup}_{x\in J}\int_a^x \mathrm{d}y=b-a;$$

$$\|T\|_2\leqslant\left(\int_a^b \mathrm{d}x\int_a^x \mathrm{d}y\right)^{1/2}=\frac{b-a}{\sqrt{2}}.$$

如果说命题2.2.3与命题2.2.2完全对应的话,那么以下结果则是积分算子所独有的.

命题 2.2.4 设 $K(x,y)$ 在 $J\times J$ 上连续,积分算子 T 定义如式(2.2.8),则 $T\in L(C(J))$,且

$$\|T\|=\sup_{x\in J}\int_a^b |K(x,y)|\mathrm{d}y. \tag{2.2.10}$$

证 任给 $u\in C(J)$,显然 $Tu\in C(J)$.令

$$k(x)=\int_a^b |K(x,y)|\mathrm{d}y\quad(x\in J),$$

则 $k(\cdot)\in C(J)$,式(2.2.10)相当于 $\|T\|=\|k\|_0$.由

$$\begin{aligned}\|Tu\|_0&=\sup_{x\in J}\left|\int_a^b K(x,y)u(y)\mathrm{d}y\right|\\ &\leqslant\sup_{x\in J}\int_a^b |K(x,y)||u(y)|\mathrm{d}y\\ &\leqslant\|k\|_0\|u\|_0\qquad(\forall u\in C(J))\end{aligned}$$

得出 $T\in L(C(J))$,且 $\|T\|\leqslant\|k\|_0$.

为证 $\|T\|\geqslant\|k\|_0$,取 $x_0\in J$,使 $\|k\|_0=k(x_0)$.令 $\varphi(y)=\operatorname{sgn}K(x_0,y)$.由 Lusin 定理([8],定理2.4.4),有 $\{u_n\}\subset C(J)$,使得 $u_n\to\varphi$, a.e.,且可设 $|u_n|\leqslant 1$.于是

$$\|k\|_0=\int_a^b K(x_0,y)\varphi(y)\mathrm{d}y$$

$$= \lim_{n}\int_a^b K(x_0,y)u_n(y)\mathrm{d}y \qquad \text{(用控制收敛定理)}$$

$$= \lim_{n} Tu_n(x_0) \leqslant \overline{\lim_{n}} \| Tu_n \|_0$$

$$\leqslant \overline{\lim_{n}} \| T \| \| u_n \|_0 \leqslant \| T \|,$$

如所要证. □

对于以上所述的积分算子,有多种途径进行拓广.首先,区间 J 可代之以更一般的集合(无穷区间、多维区域甚至一般测度空间).其次,可将积分算子看作不同空间之间的算子.此处不拟详细介绍这些拓广.

下面考虑几个具有特殊形式核的积分算子,它们在应用中是特别常见的.

若取 $K(x,y) = \varphi(x-y)$ 为核,就得到如下积分算子:

$$T_\varphi u(x) = \int_{\mathbf{R}^n} \varphi(x-y)u(y)\mathrm{d}y \quad (x \in \mathbf{R}^n). \tag{2.2.11}$$

通常将式(2.2.11)右端的积分记作 $\varphi * u$,并称它为函数 φ 与 u 的**卷积**,因而由 $T_\varphi u = \varphi * u$ 定义一**卷积算子**.算子 T_φ 显然在其有定义的集合上为线性算子,其定义域与性质则决定于 φ 的选择.以下是基本的结论.

命题 2.2.5　设 $1 \leqslant p = q/(q-1) \leqslant \infty$.

(i) 若 $\varphi \in L^1(\mathbf{R}^n)$,则 $T_\varphi \in L(L^p(\mathbf{R}^n))$,且 $\| T_\varphi \| \leqslant \| \varphi \|_1$.

(ii) 若 $\varphi \in L^p(\mathbf{R}^n)$,则 $T_\varphi \in L(L^q(\mathbf{R}^n), C_b(\mathbf{R}^n))$,且 $\| T_\varphi \| \leqslant \| \varphi \|_p$,此处 $C_b(\mathbf{R}^n) = C(\mathbf{R}^n) \cap B(\mathbf{R}^n)$,其中采用 sup 范数.

(iii) 若 $\varphi \in L^2(\mathbf{R}^n)$,则 $T_\varphi \in L(L^2(\mathbf{R}^n), C_b(\mathbf{R}^n))$,且 $\| T_\varphi \| \leqslant \| \varphi \|_2$.

证　不妨设 $1 < p = q(q-1) < \infty$, $p = 1, \infty$ 的情况证明更简单.

(i) 任给 $u \in L^p(\mathbf{R}^n)$,有

$$\| T_\varphi u \|_p^p = \int_{\mathbf{R}^n} \left| \int_{\mathbf{R}^n} \varphi(x-y)u(y)\mathrm{d}y \right|^p \mathrm{d}x$$

$$\leqslant \int_{\mathbf{R}^n} \left[\int_{\mathbf{R}^n} |\varphi(x-y)|^{1/q} |\varphi(x-y)|^{1/p} |u(y)| \mathrm{d}y \right]^p \mathrm{d}x$$

$$\leqslant \int_{\mathbf{R}^n} \left[\left(\int_{\mathbf{R}^n} |\varphi(x-y)| \mathrm{d}y \right)^{p/q} \int_{\mathbf{R}^n} |\varphi(x-y)| |u(y)|^p \mathrm{d}y \right] \mathrm{d}x$$

(用 Hölder 不等式)

$$= \| \varphi \|_1^{p/q} \int_{\mathbf{R}^n} |u(y)|^p \mathrm{d}y \int_{\mathbf{R}^n} |\varphi(x-y)| \mathrm{d}x \quad \text{(积分互换)}$$

$$=(\|\varphi\|_1\|u\|_p)^p,$$

由此得 $\|T_\varphi u\|_p \leqslant \|\varphi\|_1\|u\|_p$，从而 $T_\varphi \in L(L^p(\mathbf{R}^n))$ 且 $\|T_\varphi\| \leqslant \|\varphi\|_1$.

(ii) 取定 $u \in L^q(\mathbf{R}^n)$，首先证 $T_\varphi(u) \in C(\mathbf{R}^n)$. 任给 $x, z \in \mathbf{R}^n$，约定 $\varphi_x(y) = \varphi(x+y)$，则

$$\begin{aligned}
&|T_\varphi u(x) - T_\varphi u(z)| \\
&\leqslant \int_{\mathbf{R}^n} |\varphi(x-y) - \varphi(z-y)|\,|u(y)|\,\mathrm{d}y \\
&\leqslant \Big(\int_{\mathbf{R}^n} |\varphi(x-y) - \varphi(z-y)|^p \mathrm{d}y\Big)^{1/p} \|u\|_q \quad (\text{用 Hölder 不等式}) \\
&= \Big(\int_{\mathbf{R}^n} |\varphi(x-z+y) - \varphi(y)|^p \mathrm{d}y\Big) \|u\|_q \\
&= \|\varphi_{x-z} - \varphi\|_p \|u\|_q.
\end{aligned}$$

这结合下面将补证的引理得出：若 $x, z \in \mathbf{R}^n$，$x - z \to 0$，则 $T_\varphi u(x) - T_\varphi u(z) \to 0$. 这表明 $T_\varphi u$ 在 $\mathbf{R}^n$ 上一致连续. 再用 Hölder 不等式得

$$|T_\varphi u(x)| \leqslant \Big(\int_{\mathbf{R}^n} |\varphi(x-y)|\,\mathrm{d}y\Big)^{1/p} \|u\|_q = \|\varphi\|_p \|u\|_q,$$

可见 $\|T_\varphi u\|_0 \leqslant \|\varphi\|_p \|u\|_q$. 因此 $T_\varphi \in L(L^q(\mathbf{R}^n), C_b(\mathbf{R}^n))$，$\|T_\varphi\| \leqslant \|\varphi\|_p$.

(iii) 是(ii)的特殊情况. □

现在补证所需的引理.

引理 2.2.6 设 $\varphi \in L^p(\mathbf{R}^n)(1 \leqslant p < \infty)$，$\varphi_x(y) = \varphi(x+y)$，则当 $x \in \mathbf{R}^n$，$x \to 0$ 时 $\|\varphi_x - \varphi\|_p \to 0$.

证 不妨设 $n = 1$（$n > 1$ 的情况并无本质区别，只是记号繁些）. 首先设 $\varphi = \chi_\delta$，$\delta = (\alpha, \beta) \subset \mathbf{R}$ 是一有限区间，则直接看出

$$\|\varphi_x - \varphi\|_p^p \leqslant 2|x| \to 0 \quad (x \to 0).$$

由此又推出：若 φ 是有紧支集的阶梯函数，则引理结论成立. 现在设 $\varphi \in L^p(\mathbf{R})$. $\forall \varepsilon > 0$，取有紧支集的阶梯函数 ψ，使得 $\|\varphi - \psi\|_p < \varepsilon$（参考例 1.3.8(iv)）. 取 $\delta > 0$，使当 $x \in \mathbf{R}$，$|x| < \delta$ 时 $\|\psi_x - \psi\|_p < \varepsilon$. 于是当 $|x| < \delta$ 时有

$$\|\varphi_x-\varphi\|_p\leqslant\|\varphi_x-\psi_x\|_p+\|\psi_x-\psi\|_p+\|\psi-\varphi\|_p$$
$$<\varepsilon+\varepsilon+\varepsilon=3\varepsilon,$$

其中用到 $\|\varphi_x-\psi_x\|_p=\|\varphi-\psi\|_p$(何故?).这得出所要证. □

你注意到,如同例 1.3.9 一样,以上证明大大得益于使用基本集概念所带来的简化.

在例 1.3.9 中,我们考察了 $u\in L^1(\mathbf{R})$ 的 Fourier 变换.一般地,对任给 $u\in L^1(\mathbf{R}^n)$,定义 u 的 Fourier 变换为

$$Fu(x)=\hat{u}(x)=\int_{\mathbf{R}^n}e^{-ix\cdot y}u(y)dy\quad(x\in\mathbf{R}^n),\tag{2.2.12}$$

其中 $x\cdot y=\langle x,y\rangle$ 是 $\mathbf{R}^n$ 中的标准内积.另一方面,从本节的观点来看,由式(2.2.12)定义的 F 是一个以 $K(x,y)\triangleq e^{-ix\cdot y}$ 为核的积分算子,今指出它的某些性质.首先,容易推出(用控制收敛定理) $\hat{u}\in C(\mathbf{R}^n)$.其次,注意 $-1=e^{i\pi}$,于是

$$\begin{aligned}2|\hat{u}(x)|&=|\hat{u}(x)-e^{i\pi}\hat{u}(x)|\\&=\left|\hat{u}(x)-\int_{\mathbf{R}^n}e^{-i(x\cdot y-\pi)}u(y)dy\right|\\&=\left|\hat{u}(x)-\int_{\mathbf{R}^n}e^{-ix\cdot(y-\pi x/|x|^2)}u(y)dy\right|\\&=\left|\hat{u}(x)-\int_{\mathbf{R}^n}e^{-ix\cdot y}u(y+z)dy\right|\qquad\left(z=\frac{x}{|x|^2}\right)\\&=\left|\int_{\mathbf{R}^n}e^{-ix\cdot y}[u(y)-u(y+z)]dy\right|\\&\leqslant\|u-u_z\|_1,\end{aligned}$$

其中 $u_z(y)=u(y+z)$.令 $|x|\to\infty$,则 $z=|x|^{-2}x\to0$,从而 $\|u-u_z\|_1\to0$(用引理 2.2.6),故 $\hat{u}(x)\to0$.令

$$C_0(\mathbf{R}^n)=\{u\in C(\mathbf{R}^n):\lim_{|x|\to\infty}u(x)=0\},$$

则 $C_0(\mathbf{R}^n)$ 依 sup 范数是一个 Banach 空间(参考第一章习题 10).因显然有

$$\|\hat{u}\|_0\leqslant\int_{\mathbf{R}^n}|u(x)|dx=\|u\|_1,$$

故得出 $F\in L(L^1(\mathbf{R}^n),C_0(\mathbf{R}^n))$,且 $\|F\|\leqslant1$.若取 u 为 n 维方体 $C=[-1,$

$1]^n$ 的特征函数，则 $\|u\|_1 = 2^n$；

$$\hat{u}(x) = \prod_{j=1}^{n}\int_{-1}^{1} e^{-ix_j t}\,dt = \prod_{j=1}^{n}\frac{2\sin x_j}{x_j},$$

由此易见 $\|\hat{u}\|_0 = 2^n$，因而得出 $\|F\| \geqslant 1$. 这就得出 $\|F\| = 1$.

§ 2.2.3 微分算子

像式(2.1.10)这样简单的微分算子都不是有界算子，看来这是一个不祥之兆！确实，如同空间 $C^m(\Omega)$ 具有比 $L^p(\Omega)$ 较坏的结构一样，微分算子的性质通常不及积分算子. 对微分算子的深入考察远非此处所能进行，下面的简单讨论仅起解释概念的作用.

以下设 $k, l \in \mathbf{Z}_+, m = k + l, \alpha \in \mathbf{Z}_+^n, |\alpha| = k$.

首先设 $\Omega \subset \mathbf{R}^n$ 是一有界闭区域，则微分算子

$$\partial^\alpha : C^m(\Omega) \to C^l(\Omega) \quad u \to \partial^\alpha u, \tag{2.2.13}$$

是一线性算子. 如果在 $C^m(\Omega)$ 与 $C^l(\Omega)$ 中都使用范数 $\|\cdot\|_l$（依式(1.2.6)'），则如 § 2.1 中算子(2.2.10)一样，算子 ∂^α 是无界的. 但是如果在 $C^m(\Omega)$ 中采用范数 $\|\cdot\|_m$，则 ∂^α 是有界的：

$$\begin{aligned}\|\partial^\alpha u\|_l &= \max_{|\beta|\leqslant l}\|\partial^\beta\partial^\alpha u\|_0 \\ &\leqslant \max_{|\beta|\leqslant m}\|\partial^\beta u\|_0 = \|u\|_m \quad (\forall u \in C^m(\Omega)),\end{aligned}$$

这表明 $\|\partial^\alpha\| \leqslant 1$.

类似的结论可用到微分算子：

$$\partial^\alpha : H^{m,p}(\Omega) \to H^{l,p}(\Omega), \quad u \to \partial^\alpha u, \tag{2.2.14}$$

其中 $\Omega \subset \mathbf{R}^n$ 是一开区域，$1 \leqslant p < \infty$. 事实上，对任给 $u \in H^{m,p}(\Omega)$，有

$$\begin{aligned}\|\partial^\alpha u\|_{l,p} &= \Big(\sum_{|\beta|\leqslant l}\|\partial^\beta\partial^\alpha u\|_p^p\Big)^{1/p} \\ &\leqslant \Big(\sum_{|\beta|\leqslant m}\|\partial^\beta u\|_p^p\Big)^{1/p} = \|u\|_{m,p},\end{aligned}$$

这表明对于微分算子(2.2.14)亦有 $\|\partial^\alpha\| \leqslant 1$.

不过，从应用上考虑，真正重要的是将 ∂^α 看作从 $H^m(\Omega)$ 到同一空间 $H^m(\Omega)$ 中的线性算子. 但这样一来，∂^α 就只定义于 $H^m(\Omega)$ 的某个子空间上，且是无界算子. 在 § 3.5.3 中，将对这种意义下的微分算子作点最初步的考虑.

§2.2.4　关于算子范数的评注

在上节与本节中,我们已求得若干有界线性算子的范数.无论是迫于实际需要,还是基于追求理论上完整性的偏好,人们常常致力于准确地求出所研究的有界线性算子的范数.经验表明,这未必容易成功.读者在面对这类问题时,通常有两种方法可供选择.

(A) **利用已知的标准结果**.例如本章的命题 2.2.1～命题 2.2.4,定理 2.3.2～定理 2.3.6,以及下章的命题 3.4.2 与定理 3.4.5 等.这类结果能够提供最便捷的结论,但它们在泛函分析中不是很多,其作用是有限的.

(B) **直接法**.对于 $T\in L(X,Y)$,为求 $\|T\|$,通常依如下程序进行:$\forall x\in X$,对 $\|Tx\|$ 作出一个尽可能准确的估计 $\|Tx\|\leqslant\beta\|x\|$,从而推测 $\|T\|=\beta$.为证实这一推测,可用以下方法之一:

(i) 选取适当的 $x_0\in X$,使得 $\|x_0\|=1$,$\|Tx_0\|=\beta$,例如在例 2.1.3 中就是如此.这一方法得以可行,其前提显然是 $\|T\|=\max\limits_{\|x\|=1}\|Tx\|$,即连续函数 $f(x)=\|Tx\|$ 在单位球面上取得最大值,而在无限维空间中这未必为真,因单位球面不是紧集(参考定理 1.4.4 与 1.4.8).例如对

$$J(u)=\int_{-1}^{0}u(x)\mathrm{d}x-\int_{0}^{1}u(x)\mathrm{d}x\quad(u\in C[-1,1]),$$

就不存在 $u\in C[-1,1]$,使 $\|u\|_0=1$,$|J(u)|=\|J\|=2$.因此以上方法应用很有限.

(ii) 对任给 $\alpha<\beta$,选取适当的 $x_\alpha\in X$,使得 $\|x_\alpha\|=1$,$\|Tx_\alpha\|\geqslant\alpha$.例如,在命题 2.2.2 之证中就是如此.

(iii) 选取适当的序列 $\{x_n\}\subset X$,使得 $\|x_n\|\leqslant 1$,而 $\overline{\lim\limits_{n}}\|Tx_n\|\geqslant\beta$.例如,命题 2.2.4 之证即用此法.

毫无疑问,如上的方法(i)～(iii)成功的前提是,$\|T\|=\beta$ 是一正确的猜测,即一开始所得的估计 $\|Tx\|\leqslant\beta\|x\|$ 已经是最佳的,而这未必总能保证.

应当强调指出,不能准确地求出 $\|T\|$,未必总是一个很严重的问题.例如,对于命题 2.2.2(iii)中的 A,我们只能断定 $\|A\|_2\leqslant\beta$(有例子指出,不严格不等号是可能的),但这并不妨碍我们有效地运用算子 A.在很多情况下,我们只要能得出关于 $\|T\|$ 的某个合乎需要的估计(例如 $\|T\|<1$),或者甚至只要知道 $\|T\|<\infty$,或许就够了.

§2.3　对偶空间与对偶算子

§2.3.1　有界线性泛函

给定 $\mathbf{K}$ 上的赋范空间 X，约定 $X^* = L(X,\mathbf{K})$. 由命题2.1.6, X^* 是 $\mathbf{K}$ 上的 Banach 空间，称为 X 的**对偶空间**；称每个 $u \in X^*$ 为 X 上的**有界线性泛函**.

因有界线性泛函是有界线性算子的特殊情况，故关于一般有界线性算子的概念与结论，必定亦适用于有界线性泛函. 因此，§2.1 中的结果经适当表述之后，就成为关于有界线性泛函的相应结论. 例如，对任给 $u \in X^*$，直接由 §2.1 中的式(2.1.5)～(2.1.9)有

$$\begin{aligned}\|u\| &= \sup_{x\neq 0} |u(x)|/\|x\| \\ &= \sup_{\|x\|=1} |u(x)| \\ &= \sup_{\|x\|\leqslant 1} |u(x)| \\ &= \min\{k \geqslant 0 : |u(x)| \leqslant k\|x\| (\forall x \in X)\};\end{aligned} \tag{2.3.1}$$

$$|u(x)| \leqslant \|u\| \|x\| \tag{2.3.2}$$

对于 X 上的线性泛函 u, u 有界$\Leftrightarrow$ u 连续. 其他结论不必一一写出. 不过，鉴于有界线性泛函的特殊性，我们更关注那些对于它所独有的结论.

有界线性泛函的一个突出优点是它有明显的几何意义. 如你熟知，$\mathbf{R}^3$ 中任何平面可表为线性方程

$$\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 = c,$$

其中 $0 \neq (\alpha_1,\alpha_2,\alpha_3)^{\mathrm{T}} \in \mathbf{R}^3$. 鉴于此，对于任何 $0 \neq f \in X^*$ 与 $c \in \mathbf{K}$，称

$$f^{-1}(c) = \{x \in X : f(x) = c\}$$

为 X 中由 f 决定的**超平面**，也记作 $\{f = c\}$. 特别，过原点的超平面 $f^{-1}(0) = N(f)$ 是 X 的闭子空间(依命题 2.1.1(i)与定理 1.3.6). 任何超平面 $f^{-1}(c)$ 必可由 $N(f)$ 平移得到：任取 $x_0 \in f^{-1}(c)$，则易验证 $f^{-1}(c) = x_0 + N(f)$. 超平面族 $\{x + N(f) : x \in X\}$ 填满了整个空间 X，其中任意两个超平面可看作互相平行. 那么，$N(f)$ 以什么特点与 X 的其他子空间相区别呢？这由以下结果解答.

命题 2.3.1　设 $A \subset X$ 是一子空间. 则以下两条件等价：

(i) 有 $0 \neq f \in X^*$，使得 $A = N(f)$；除一个常数因子的差别外，f 由 A 惟一决定.

(ii) 存在拓扑直和分解 $X = A \oplus \mathbf{K}x_0$，此处 $x_0 \neq 0$，$\mathbf{K}x_0 = \{\lambda x_0 : \lambda \in \mathbf{K}\}$ 是 X 中的由 x_0 生成的 1 维子空间(可理解为过点 0 与 x_0 的直线，见图 2-1).

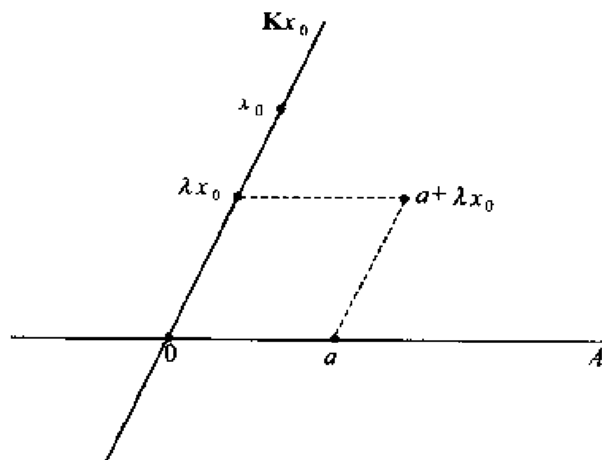


图 2-1

证　(i)⇒(ii). 设 $A = N(f)$，$0 \neq f \in X^*$，$f(x_0) \neq 0$. 则每个 $x \in X$ 可分解为

$$x = \left[x - \frac{f(x)}{f(x_0)}x_0\right] + \frac{f(x)}{f(x_0)}x_0,$$

上式右端两项分别属于 A 与 $\mathbf{K}x_0$. 若 $\lambda x_0 \in A$，则由 $0 = f(\lambda x_0) = \lambda f(x_0)$ 推出 $\lambda = 0$，这表明 $A \cap \mathbf{K}x_0 = \{0\}$. 因此 $X = A \oplus \mathbf{K}x_0$. 因 A 与 $\mathbf{K}x_0$ 均为 X 的闭子空间，故 $A \oplus \mathbf{K}x_0$ 是拓扑直和.

(ii)⇒(i). 设有拓扑直和分解 $X = A \oplus \mathbf{K}x_0$. 因每个 $x \in X$ 有惟一分解 $x = a + \lambda x_0$，$a \in A$，$\lambda \in \mathbf{K}$，定义 $f(x) = \lambda$，则 $f(x)$ 由 x 惟一决定. 于是得到一个泛函 $f: X \to \mathbf{K}$，直接看出它是线性的. 显然 $f(x) = 0 \Leftrightarrow x \in A$，故 $A = N(f)$. 若 f 无界，则必有序列 $\{x_n\} \subset X$，使得 $|f(x_n)| > n\|x_n\|$. 任给 $x \in X$，有

$a_n \triangleq x_0 - \dfrac{f(x_0)}{f(x_n)}x_n \in A$，

$$\lim_n a_n = x_0 - \lim_n \frac{f(x_0)}{n} \cdot \frac{nx_n}{f(x_n)} = x_0.$$

这表明 $x_0 \in \bar{A} - A$，得出矛盾. 因此 f 有界，即 $f \in X^*$. 余下指出：除常数因子的差别外，f 由 A 惟一决定. 若 $g \in X^*$ 满足 $N(g) = A$，则 $\forall x \in X$，有

$$g(x) = g(x - f(x)x_0) + g(f(x)x_0)$$

$$= g(f(x)x_0) = f(x)g(x_0) = cf(x),$$

其中 $c = g(x_0)$ 与 x 无关. □

由命题 2.3.1 得出,若 $0 \neq f \in X^*$, $f(x_0) \neq 0$, 则

$$X = N(f) \oplus \mathbf{K}x_0 \tag{3}$$

这一分解对有界线性泛函的研究是基本的.

§2.3.2 表示定理

在逻辑上, X^* 是一个完全确定的 Banach 空间,它的元素就是 X 上的有界线性泛函,似乎不存在"求出" X^* 的问题.不过,若 X^* 缺乏某种直观形象,以至难以有效地思考与运用,那么,我们实际上会把它当成一种未知的东西.这就提出一个问题:如何将 X^* **具体**表示出来? 这正是表示定理所要解决的课题.

表示问题的一般思路如下:对于给定的赋范空间 X, 确定一个 Banach 空间 Y, 它通常是已被充分研究因而相当熟悉的空间,使得存在等距同构

$$T: Y \to X^*, \quad y \to \varphi_y. \tag{2.3.4}$$

因而由式(2.3.4)得出结论: $u \in X^*$ 有**通式**

$$u(x) = \varphi_y(x) \quad (x \in X),$$

其中 $y \in Y$ 由 u 惟一决定,且 $\|y\| = \|u\|$. 若将 $u = \varphi_y$ 与 y 视为等同,则不妨认定 $X^* = Y$. 这样,通过同构对应式(2.3.4),本来很**抽象**的空间 X^* 就获得了一种**具体**的表示, Y 就是 X^* 的一个**表示**,或称为一个**实现**.若令 $\varphi(x,y) = \varphi_y(x)$, φ_y 依式(2.3.4),则它是双线性的(即分别对 x 与 y 是线性的),且满足不等式

$$|\varphi(x,y)| \leqslant \|\varphi_y\| \|x\| = \|x\| \|y\|. \tag{2.3.5}$$

可见, $\varphi(x,y)$ 显示出内积的某些特征(参照定义 1.5.1),例如式(2.3.5)恰与 Schwarz 不等式相对应.这启示我们利用类似于内积的表达式去构成所需的函数 φ. 下面就依照上述思路来建立若干表示定理.

定理 2.3.2 设 $1 \leqslant p = q/(q-1) < \infty$, 则 $(l^p)^* \cong l^q$($\cong$ 表等距同构,本节中概如此)①; $u \in (l^p)^*$ 有通式

---

① 使用记号 $X^* \cong Y$ 时,总应将$\cong$理解为由某个表示定理所确定的同构;仅当这样的同构存在时,才可认定 Y 为 X 的对偶空间.

$$u(x)=\sum_i x_i y_i \quad (x=(x_i)\in l^p), \tag{2.3.6}$$

其中 $y=(y_i)\in l^q$ 由 u 惟一决定，且 $\|y\|_q=\|u\|$.

证　定义 $\varphi: l^p\times l^q\to \mathbf{K}$ 如下：

$$\varphi(x,y)=\sum_i x_i y_i \quad (x=(x_i)\in l^p, y=(y_i)\in l^q). \tag{2.3.7}$$

显然 $\varphi(x,y)$ 是双线性的，且 $|\varphi(x,y)|\leqslant \|x\|_p\|y\|_q$ (用式(1.2.10)). 这就表明，若令 $\varphi_y=\varphi(\cdot,y)$，则 $\varphi_y\in(l^p)^*$ 且 $\|\varphi_y\|\leqslant\|y\|_q$. 于是

$$T: l^q\to(l^p)^*, \quad y\to\varphi_y$$

是一线性算子，它满足 $\|Ty\|\leqslant\|y\|_q$. 余下只要对任给 $u\in(l^p)^*$，证明以下两件事：

(i) 存在 $y\in l^q$，使得 $u=Ty$，这意味着 u 能表如式(2.3.6)；

(ii) $\|y\|_q\leqslant\|u\|$，

这将推出 T 是满射且 $\|Ty\|=\|y\|_q$，从而 T 为等距同构.

取定 $u\in(l^p)^*$. 设 $\{e_i\}$ 是 l^p 的标准基，令 $y_i=u(e_i)$，$y=(y_i)$. $\forall x=(x_i)\in l^p$，有

$$u(x)=u\left(\lim_n\sum_{i=1}^n x_i e_i\right)=\lim_n\sum_{i=1}^n x_i u(e_i)$$

$$=\lim_n\sum_{i=1}^n x_i y_i=\sum_{i=1}^\infty x_i y_i,$$

可见式(2.3.6)成立. 若 $p=1$，则 $q=\infty$，

$$\|y\|_\infty=\sup_i|y_i|=\sup_i|u(e_i)|\leqslant\|u\|.$$

余下只要对 $1<p<\infty$ 证 $\|y\|_q\leqslant\|u\|$；为此，又只要对任何 $n\geqslant 1$ 证 $\left(\sum_1^n|y_i|^q\right)^{1/q}\leqslant\|u\|$，这由以下推演得出：

$$\sum_{i=1}^n|y_i|^q=\sum_{i=1}^n|y_i|^{q-1}y_i\,\mathrm{sgn}\,y_i{}^{①}$$

$$=\sum_{i=1}^n|y_i|^{q-1}(\mathrm{sgn}\,y_i)u(e_i)$$

$$=u\left(\sum_{i=1}^n|y_i|^{q-1}(\mathrm{sgn}\,y_i)e_i\right)$$

① 参看 70 页注①.

$$\leqslant \|u\| \left\| \sum_{i=1}^{n} |y_i|^{q-1}(\mathrm{sgn} y_i) e_i \right\|_p$$

$$= \|u\| \left(\sum_{i=1}^{n} |y_i|^{p(q-1)} \right)^{1/p}$$

$$= \|u\| \left(\sum_{i=1}^{n} |y_i|^{q} \right)^{1/p}, \qquad (\text{用 } p(q-1) = q)$$

以上不等式两端同时除以 $\left(\sum_1^n |y_i|^q\right)^{1/p}$，即得所要证. □

至此，你大概已猜测到 $(L^p)^* \cong L^q$，事情的确如此.

定理 2.3.3 设 $1 \leqslant p = q/(q-1) < \infty, J = [a,b](a < b)$，则 $L^p(J)^* \simeq L^q(J)$；$f \in L^p(J)^*$ 有通式

$$f(u) = \int_a^b u(x)v(x)\mathrm{d}x \quad (u \in L^p(J)), \tag{2.3.8}$$

其中 $v \in L^q(J)$ 由 f 惟一决定，且 $\|v\|_q = \|f\|$.

证明的思路大体上类似于定理 2.3.2 之证，对应于式(2.3.7)，考虑双线性泛函

$$\varphi(u,v) = \int_a^b u(x)v(x)\mathrm{d}x \quad (u \in L^p(J), v \in L^q(J)).$$

不过，这次将遇到某些技术性困难，因而证明要繁琐些，只得从略(参考文献[10]§2.10).

其次要指出，将 J 换成更一般的 Ω (加上某些限制)之后，定理 2.3.3 的结论依然成立. 这样，对于§1.2 中所述的空间系列 $L^p(1 \leqslant p < \infty)$，求对偶空间的问题就得到了解决，而且对偶空间就在同一空间系列之内. 尤其值得注意的是，若 $1 < p < \infty$，则空间 L^p 与 L^q 互为对偶空间，此处 $q = p/(p-1)$；若对空间 L^p 取两次对偶，则仍得到 L^p. 一般地，令 $X^{**} = (X^*)^*$，称它为 X 的**二次对偶**. 若 $X^{**} \simeq X$，则称 X 为**自反空间**(更准确的定义见定义 2.4.11). 因此，$L^p(1 < p < \infty)$ 就是自反空间. 正是这一事实使得空间 L^p 具有明显优势.

至于§1.2.3 中所述的另一空间系列 $C^m(\Omega)$，就没有如 L^p 这样好的结果. 即使描述最简单的空间 $C(J)$ 的对偶空间，也需要用到一个新的函数类：所谓有界变差函数类. 下面不加证明地给出一个结果.

定理 2.3.4 设 $J = [a,b](a < b)$，则 $C(J)^* \simeq \mathrm{BV}_0(J)$，其中

$$\mathrm{BV}_0(J) = \{v : v \text{ 在}[a,b]\text{ 上有界变差、右连续且 } v(a) = 0\};$$

在 $\mathrm{BV}_0(J)$ 中使用范数 $\|v\| = \mathrm{V}_a^b(v)$. $f \in C(J)^*$ 有通式

$$f(u)=\int_a^b u(x)\mathrm{d}v(x)\quad(u\in C(J)),\tag{2.3.9}$$

其中 $v\in \mathrm{BV}_0(J)$ 由 f 惟一决定，且 $\|v\|=\|f\|$，式(2.3.9)中的积分是所谓 Riemann-Stieltjes 积分.

特别，任给 $v\in L^1(J)$，令 $w(x)=\int_a^x v(t)\mathrm{d}t$，则必 $w\in \mathrm{BV}_0(J)$，且可验证

$$\mathrm{V}_a^b(w)=\int_a^b |v(x)|\mathrm{d}x=\|v\|_1;$$

$$\int_a^b u(x)\mathrm{d}w(x)=\int_a^b u(x)v(x)\mathrm{d}x\quad(u\in C(J)).$$

因此，由

$$f(u)=\int_a^b u(x)v(x)\mathrm{d}x\qquad(u\in C(J)),\tag{2.3.10}$$

定义出一个 $f\in C(J)^*$，且 $\|f\|=\|v\|_1$. 于是可以说，$L^1(J)$ 可等距嵌入到 $C(J)^*$ 中(参考定理 1.1.6).

现在用简单例子来说明如何应用定理 2.3.3~2.3.4.

例 2.3.5　(i) 设 $1<p=q/(q-1)<\infty, 0<\alpha<p$,

$$f(u)=\int_0^1 u(x^\alpha)\mathrm{d}x,\quad u\in L^p[0,1].$$

今指明 $f\in L^p[0,1]^*$. 首先对积分作变量代换 $y=x^\alpha$:

$$f(u)=\frac{1}{\alpha}\int_0^1 y^{\frac{1}{\alpha}-1}u(y)\mathrm{d}y.$$

令 $v(y)=\frac{1}{\alpha}y^{\frac{1}{\alpha}-1}$，则

$$\begin{aligned}\|v\|_q&=\frac{1}{\alpha}\left(\int_0^1 y^{q\left(\frac{1}{\alpha}-1\right)}\mathrm{d}y\right)^{1/q}\\&=\frac{1}{\alpha}\left[\frac{1}{q(\alpha^{-1}-1)+1}\right]^{1/q}\\&=\frac{1}{\alpha}\left[\frac{\alpha}{(p-\alpha)(q-1)}\right]^{1/q}<\infty.\quad(\text{用 }0<\alpha<p)\end{aligned}$$

于是由定理 2.3.3 有 $f\in L^p[0,1]^*$ 且 $\|f\|=\|v\|_q$. 你不妨试探一下，如果不用定理 2.3.3，是否容易得出此结果?

(ii) 设 $J=[0,1], 0<a<b<1, \alpha,\beta\in\mathbf{K}$，定义

$$f(u) = \alpha u(a) + \beta u(b) \quad (u \in C(J)).$$

今用定理 2.3.4 说明 $f \in C(J)^*$ 并求出 $\|f\|$. 为此,作函数

$$v(x) = \begin{cases} 0, & 0 \leqslant x < a; \\ \alpha, & a \leqslant x < b; \\ \alpha + \beta, & b \leqslant x \leqslant 1. \end{cases}$$

则 $v \in BV_0(J), V_0^1(v) = |\alpha| + |\beta|$, 且

$$f(u) = \int_0^1 u(x) \mathrm{d}v(x) \quad (u \in C(J)).$$

于是由定理 2.3.4 有 $f \in C(J)^*$, $\|f\| = V_0^1(v) = |\alpha| + |\beta|$.

现在再回到定理2.3.3.取 $p = q = 2$ 得到 $(L^2)^* \cong L^2$,这表明空间 L^2 是**自对偶**的.我们猜想一般 Hilbert 空间亦有此性质,事情确实如此.

定理 2.3.6 设 H 是实 Hilbert 空间,则 $H^* \simeq H, u \in H^*$ 有通式

$$u(x) = \langle x, y \rangle \quad (x \in H), \tag{2.3.11}$$

其中 $y \in H$ 由 u 惟一决定,且 $\|y\| = \|u\|$.

证 令 $\varphi(x,y) = \langle x, y \rangle (x, y \in H)$, 则如同定理 2.3.2 之证,可指明

$$T: H \to H^*, \quad y \mapsto \varphi_y = \langle \cdot, y \rangle \tag{2.3.12}$$

是一个线性算子,且 $\|Ty\| \leqslant \|y\|$. 余下只要证明:任给 $u \in H^*$,有 $y \in H$,使得 $u(x)$ 表示如式(11),且 $\|y\| \leqslant \|u\|$. 可设 $u \neq 0$(否则取 $y = 0$ 好了),则 $A = N(u)$ 是 H 的真闭子空间.由正交分解定理1.5.10,有 $H = A \oplus A^{\perp}$. 取 $x_0 \in A^{\perp}$, 使 $u(x_0) \neq 0$(此种 x_0 必存在,何故?),不妨设 $u(x_0) = 1$. 令 $y = x_0 / \|x_0\|^2$, 则 $y \in A^{\perp}$. 直接看出 $x - u(x)x_0 \in A$, 于是

$$\begin{aligned} \langle x, y \rangle &= \langle x - u(x)x_0, y \rangle + \langle u(x)x_0, y \rangle \\ &= u(x)\langle x_0, y \rangle = u(x) \qquad (x \in H), \end{aligned}$$

这得出式(2.3.11).取 $x = y$ 得 $\|y\|^2 = u(y) \leqslant \|u\| \|y\|$, 于是 $\|y\| \leqslant \|u\|$, 如所要证. □

式(2.3.11)称为空间 H 上有界线性泛函的 **Riesz 表示**,它是 Hilbert 空间理论中最有价值的结果之一.若 H 是复 Hilbert 空间,则 Riesz 表示仍然适用,不同的只是,式(2.3.12)是一等距的共轭线性同构,而不是线性同构.因此,对于实或复 Hilbert 空间 H, 都可认定 $H^* = H$, 即 Hilbert 空间有自对偶性. Hilbert 空间上的有界线性泛函有 Riesz 表示这一极简单的通式,实在是一件可庆幸的事.这一事实在

算子理论中影响深远(例如参看§3.4, §3.5),此处姑且不论;一个最直接的结果是,在 Hilbert 空间中运用有界线性泛函成为一件相对简单的事. 与之相比较,在一般的 Banach 空间中却无类似的结果可用.

Riesz 表示式(2.3.11)虽然仅适用于 Hilbert 空间,但它所激发的启示亦有助于对赋范空间的研究. 为便于类比,将任何赋范空间 X 上的有界线性泛函表为

$$u(x)=\langle u,x\rangle \quad (u\in X^*, x\in X), \tag{2.3.13}$$

这就使它具有某种"内积"的外观①. 表达式(2.3.13)的好处是,其中 u 与 x 处于对等地位,在需要 u 与 x 互换角色的问题中有其方便. $\langle u,x\rangle$ 虽然一般不是内积,但它确具有内积的一些特征. 例如 $\langle u,x\rangle$ 是双线性的,而不等式

$$|\langle u,x\rangle|\leqslant \|u\|\,\|x\| \qquad (\text{依式}(2.3.2))$$

则正相当于 Schwarz 不等式. 为加强这一类比,进一步约定一些形式记号与术语:若 $\langle u,x\rangle=0$, 则说 u 与 x **正交**,并记作 $u\perp x$. 对于 $A\subset X$ 与 $U\subset X^*$, 令(参照定义 1.5.4(iii))

$$\begin{cases} A^{\perp}=\{u\in X^*: u\perp a(\forall a\in A)\}, \\ {}^{\perp}U=\{x\in X: u\perp x(\forall u\in U)\}, \end{cases} \tag{2.3.14}$$

二者分别称为 A 与 U 的**零化子**. 如果 X 是 Hilbert 空间,认定 $X=X^*$, 则由(2.3.14)定义的 $A^{\perp}$ 正好就是 A 的正交补.

§2.3.3　对偶算子

对偶空间的重要性在于, X 与 X^* 两者的性质之间存在着很强的联系,正是这种联系,使得 X^* 成为研究空间 X 的重要工具. 实际上,这种对偶关系诱导出有界线性算子与其对偶算子之间的对偶关系,这正是现在就要考虑的.

定义 2.3.7　设 $T\in L(X,Y)$, 称算子

$$T^*: Y^*\to X^*, \quad v\to v\circ T \tag{2.3.15}$$

为 T 的**对偶算子**.

必须说明一下,式(2.3.15)的确合理地定义一个算子 $T^*: Y^*\to X^*$, 且 T^* 是有界线性算子. $\forall v\in Y^*$, 显然 $T^*v=v\circ T\in X^*$. 其次, $v\circ T$ 显然对 v 是线性的. 由

$$\|T^*\|=\sup_{\|v\|\leqslant 1}\|T^*v\|=\sup_{\|v\|\leqslant 1}\sup_{\|x\|\leqslant 1}|v(Tx)|$$

① 确有作者称 $\langle u,x\rangle$ 为 u 与 x 的内积,本书不用这一名称.

$$\leqslant \sup_{\|v\|\leqslant 1, \|x\|\leqslant 1} \|v\| \|T\| \|x\| \leqslant \|T\|$$

得出 $T^* \in L(Y^*, X^*)$，且 $\|T^*\| \leqslant \|T\|$.

利用等式(2.3.13)，可将 T^* 的定义式(2.3.15)改写成

$$\langle T^* v, x\rangle = \langle v, Tx\rangle \quad (v \in Y^*, x \in X). \tag{2.3.16}$$

在对偶算子理论中，将反复用到恒等式(2.3.16).

下面的例子将为你提供关于对偶算子的某些直观印象.

例 2.3.8 (i) 设 $A = [a_{ij}] \in \mathbf{R}^{m\times n}$. $\forall x = (x_j) \in \mathbf{R}^n, y = (y_i) \in \mathbf{R}^m$，依式(2.3.16)有

$$\langle A^* y, x\rangle = \langle y, Ax\rangle = \sum_i y_i \sum_j a_{ij} x_j$$

$$= \sum_j x_j \sum_i a_{ij} y_i,$$

这表明 $(A^* y)_j = \sum_i a_{ij} y_i$，可见 $A^* = A^{\mathrm{T}}$ (A 的转置). 于是可以说，对偶算子的有限维原型，原来就是转置矩阵.

(ii) 设 $T \in L(L^2(J))$ 依式(2.2.8)，$\int_a^b \int_a^b |K(x,y)|^2 \mathrm{d}x\mathrm{d}y < \infty$，则 $T^* \in L(L^2(J))$，$\forall u, v \in L^2(J)$，依式(2.3.16)有

$$\langle T^* v, u\rangle = \langle v, Tu\rangle = \int_a^b v(x)\mathrm{d}x \int_a^b K(x,y)u(y)\mathrm{d}y$$

$$= \int_a^b u(y)\mathrm{d}y \int_a^b K(x,y)v(x)\mathrm{d}x$$

$$= \int_a^b u(x)\mathrm{d}x \int_a^b K(y,x)v(y)\mathrm{d}y,$$

故

$$T^* v(x) = \int_a^b K(y,x)v(y)\mathrm{d}y.$$

可见，T^* 是以 $K(y,x)$ 为核的积分算子，它的核 $K(y,x)$，可看作算子 T 的核 $K(x,y)$ 的“转置”.

将以上实例放在心上，下面的结果就很好理解了.

命题 2.3.9 对 $T, S \in L(X,Y), A \in L(Y,Z), \alpha, \beta \in \mathbf{K}$，以下结论成立：

$$(\alpha T + \beta S)^* = \alpha T^* + \beta S^*; \tag{2.3.17}$$

$$(AT)^* = T^* A^*; \tag{2.3.18}$$

$$(T^{-1})^* = (T^*)^{-1};\quad (\text{若 } T \text{ 是拓扑同构}) \tag{2.3.19}$$

$$\|T^*\| = \|T\|. \tag{2.3.20}$$

证　利用恒等式(2.3.16)易直接验证等式(2.3.17)与(2.3.18).若 T 为拓扑同构,则将式(2.3.18)用到等式(参看式(2.1.4))

$$T^{-1}T = I_X \text{ 与 } TT^{-1} = I_Y$$

得出：$T^*(T^{-1})^*$ 与 $(T^{-1})^*T^*$ 分别为 X^* 与 Y^* 上的单位算子,这得出式(2.3.19).至于等式(2.3.20),前面已经指出 $\|T^*\| \leqslant \|T\|$,相反的不等式将在后面补证. □

§2.4　基本定理

有几个著名定理:逆算子定理、一致有界原理与 Hahn-Banach 定理,其应用贯穿于整个泛函分析理论的展开,因而被恰当地称为**三大基本定理**.这些定理的结论从直观上看来远非一目了然,因而具有异乎寻常的深刻性.本节着重阐明这些定理的价值与一般用法,并尽可能给出其证明的容易部分,而完全的证明则请参考有关书籍.

§2.4.1　逆算子定理

设 $T: X \to Y$ 是一个线性同构.为判定 T 为拓扑同构,你必须做两件事:判定 T 与 T^{-1} 都连续.不过,若 X 与 Y 是 Banach 空间,则你仅有一半的事情要做:判定了 T(或 T^{-1})连续,即可断定 T 为拓扑同构.这就是如下著名定理的结论.

定理 2.4.1(逆算子定理)　设 X 与 Y 是 Banach 空间,$T \in L(X, Y)$ 是双射,则 T^{-1} 连续.换言之,**Banach 空间之间的连续线性同构必为拓扑同构**.

定理的证明基于 Baire 纲定理(1.4.11),颇不简单,欲知其详者可参看文献[10]§2.10.

若只假定 X 完备,则当 $T: X \to Y$ 是拓扑同构时,Y 亦必完备[①].因此逆算子定理表明,若 X 完备,$T \in L(X, Y)$ 为双射,则 T^{-1} 连续 $\Leftrightarrow Y$ 完备.这就说明了,在逆算子定理中,X 与 Y 皆完备这一条件是不可缺少的.由逆算子定理也推出:若 X 与 Y 均完备,$T \in L(X, Y)$,$N(T) = \{0\}$,$R(T)$ 是 Y 的闭子空间,则 T 是 Banach 空间 X 与 $R(T)$ 之间的拓扑同构,从而 $T^{-1} \in L(R(T), X)$.在很多情况

① 若 $\{y_n\} \subset Y$ 是 Cauchy 列,则由 T^{-1} 有界推出 $\{T^{-1}y_n\}$ 是 X 中的 Cauchy 列;设 $T^{-1}y_n \to x \in X$,则 $y_n \to Tx\ (n \to \infty)$.

下,就是以这种方式应用逆算子定理来得出 T 有有界逆.

试看一具体例子.设 $J=[a,b](a<b)$,令

$$C_0^1(J)=\{u\in C^1(J): u(a)=0\};$$

在 $C(J)$ 与 $C_0^1(J)$ 中均使用 sup 范数.则

$$T: C(J)\to C_0^1(J), u(x)\to\int_a^x u(y)\mathrm{d}y \tag{2.4.1}$$

是一线性同构(请验证).因

$$T^{-1}=\frac{\mathrm{d}}{\mathrm{d}x}: C_0^1(J)\to C(J),\quad u\to u'$$

是无界的(参考例 2.1.5,研究 $C_0^1(J)$ 中的序列 $u_n(x)=\sin n(x-a)$),故由逆算子定理知 $C_0^1(J)$ 必非 $C(J)$ 的闭子空间.

作为逆算子定理的另一个应用,下面证明所谓闭图像定理,后者亦有重要应用价值.

定理 2.4.2(闭图像定理) 设 X,Y 是 Banach 空间.若线性算子 $T: X\to Y$ 满足条件:

$$\left.\begin{array}{l}x_n\to x\\ Tx_n\to y\end{array}\right\}\Rightarrow y=Tx, \tag{2.4.2}$$

则 T 连续.

条件(2.4.2)式相当于:T 的**图像**

$$G\triangleq\{(x,Tx): x\in X\}$$

是积空间 $X\times Y$ 中的闭集.当此条件满足时,称 T 为**闭线性算子**.于是将定理 2.4.2 重新表述为:**Banach 空间之间的闭线性算子必为有界算子**。

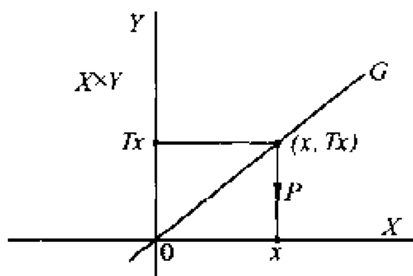


图 2-2

证 在定理条件下,G 是 Banach 空间 $X\times Y$(参看第一章习题 4)的闭子空间,因而亦是 Banach 空间.定义算子(见图 2-2)

$$P: G\to X,\quad (x,Tx)\to x,$$

则直接看出 P 是一个线性同构.回忆一下 §1.1.3 中已指出,积空间中的收敛归结为依坐标收敛.因此,对任给序列 $\{x_n\}\subset X$,有

$$(x_n,Tx_n)\to(x,Tx)\Rightarrow x_n\to x\quad(n\to\infty),$$

故 P 连续. 由定理 2.4.1, P^{-1} 亦应连续, 这意味着

$$x_n \to x \Rightarrow (x_n, Tx_n) \to (x, Tx) \quad (n \to \infty),$$

于是 $x_n \to x \Rightarrow Tx_n \to Tx (n \to \infty)$, 这正表示 T 连续. □

§2.4.2　一致有界原理

以下的定理涉及一定的算子族 $\{T_i : i \in I\} \subset L(X, Y)$, 此处 I 是任何指标集(不必是可数集). 若 $\{T_i\}$ 依算子范数有界, 即 $\sup\limits_i \| T_i \| < \infty$, 则说算子族 $\{T_i\}$ **一致有界**. 经常需要判定算子族的一致有界性, 以下定理表明, 一致有界性可从形式上弱得多的条件推出.

定理 2.4.3(一致有界原理)　设 $\{T_i : i \in I\} \subset L(X, Y)$,

$$A = \{x \in X : \sup_i \| T_i x \| < \infty\}. \tag{2.4.3}$$

(i) 若 A 是 X 中的第二纲集, 则 $\{T_i\}$ 一致有界.

(ii) 若 X 完备, $\forall x \in X$, 有 $\sup\limits_i \| T_i x \| < \infty$ (这相当于 $A = X$), 则 $\{T_i\}$ 一致有界.

(iii) 若 X 完备, $\sup\limits_i \| T_i \| = \infty$, 则 A 是 X 中的第一纲集, 从而 A^c 是第二纲集与稠集.

(iv) 若 X 完备, $F \subset X^*$, 则 F 在 X^* 中有界 $\Leftrightarrow \forall x \in X, F(x) \triangleq \{u(x) : u \in F\}$ 在 $\mathbf{K}$ 中有界.

其中(iii)的结论可形象地表述为: 对几乎所有的 $x \in X$, 成立

$$\sup_i \| T_i x \| = \infty. \tag{2.4.4}$$

满足条件(2.4.4)式的 x 称为**共鸣点**, 因而定理 2.4.3 有**共鸣定理**之称.

证　结论(ii), (iii)是(i)与 Baire 纲定理的推论. 取 $Y = \mathbf{K}, \{T_i\} = F$, 则从(ii)推出(iv). 因此只要证结论(i). 令

$$A_n = \{x \in X : \sup_i \| T_i x \| \leqslant n\}, \quad n = 1, 2, \cdots.$$

注意 $\| T_i x \| \leqslant n$ 相当于 $x \in T_i^{-1} \bar{B}_n(0)$, 由此知

$$A_n = \bigcap_i T_i^{-1} \bar{B}_n(0), \quad T_i A_n \subset \bar{B}_n(0).$$

因此 A_n 是闭集(用定理 1.3.6 与命题 1.3.4). 显然 $A = \bigcup_1^\infty A_n$. 因 A 不是第一纲集, 故有某个 A_n 含内点. 固定此 A_n, 并设 $\bar{B}_r(a) \subset A_n, r > 0$. 任给 $x \in X$, $\| x \| = 1, i \in I$, 今证 $\| T_i x \| \leqslant 2n/r$ (倘如此则有 $\sup\limits_i \| T_i \| \leqslant 2n/r < \infty$). 这由以下推导得出:

$$\| T_i x \| = \frac{1}{r} \| T_i(a + rx) - T_i a \|$$

$$\leqslant \frac{1}{r} [\| T_i(a + rx) \| + \| T_i a \|]$$

$$\leqslant 2n/r. \quad (\text{用 } T_i \bar{B}_r(a) \subset T_i A_n \subset \bar{B}_n(0)) \qquad \square$$

下面着重解释一下如何应用定理 2.4.3. 定理 2.4.3 或许是泛函分析中最常用的定理之一,其应用典型地循两条途径. 其一是应用由定理 2.4.3 导出的某些标准结果,这多半用来解决判定有界性的问题;其二就是直接应用定理 2.4.3,在很多情况下是应用其中的结论(iii). 用法必然是多种多样的,若要勉强归纳出某种模式,则可表述为:为判明某个性质 P 在 Banach 空间 X 中具有一般性,构造一个与之相关的算子序列 $\{T_n\} \subset L(X, Y)$, 使之满足 $\sup_n \| T_n \| = \infty$, 且 $\sup_n \| T_n x \| = \infty \Rightarrow x$ 满足 P. 所用的技巧依问题的具体特性而异. 下面是一个有趣的例子.

例 2.4.4 几乎每个连续函数的 Fourier 级数在一稠集上发散.

证 令 $J = [-\pi, \pi]$. 首先取定 $x \in J$. 任给 $u \in C(J)$, 以 $S_n u$ 记 u 的 Fourier 级数的部分和,并令 $T_n u = S_n u(x)$, 则

$$\sup_n | T_n u | = \infty \Rightarrow u \text{ 的 Fourier 级数在 } x \text{ 发散}.$$

在微积分学中已导出公式

$$T_n u = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\sin\left(n + \frac{1}{2}\right)(x - y)}{\sin \frac{x - y}{2}} u(y) \mathrm{d}y.$$

对照式(2.3.10)得出 $T_n \in C(J)^*$, 且

$$\| T_n \| = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{\sin\left(n + \frac{1}{2}\right)(x - y)}{\sin \frac{x - y}{2}} \right| \mathrm{d}y$$

$$= \frac{1}{\pi} \int_0^{\pi} \frac{| \sin(2n + 1) y |}{\sin y} \mathrm{d}y \qquad \left(\text{以 } y \text{ 代} \frac{x - y}{2}\right)$$

$$\geqslant \frac{1}{\pi} \int_0^{\pi} \frac{| \sin(2n + 1) y |}{y} \mathrm{d}y \qquad (\text{用 } \sin y \leqslant y)$$

$$= \frac{1}{\pi} \int_0^{(2n+1)\pi} \frac{| \sin y |}{y} \mathrm{d}y. \qquad (\text{用 } y \text{ 代} (2n + 1) y)$$

因积分 $\int_0^{\infty} \frac{| \sin y |}{y} \mathrm{d}y$ 发散,故 $\sup_n \| T_n \| = \infty$. 于是由定理 2.4.3(iii)推出,

$$A_x \triangleq \left\{ u \in C(J) : \sup_n | S_n u(x) | < \infty \right\}$$

是 $C(J)$ 中的第一纲集. 取 J 中任一可数稠子集 $\{x_i\}$ (例如取 J 中的有理点集), 则 $A = \bigcup_i A_{x_i}$ 是 $C(J)$ 中的第一纲集. 于是由 Baire 纲定理得出, A^c 是 $C(J)$ 中的第二纲集与稠集. 注意

$$A^c = \left\{ u \in C(J) : \sup_n | S_n u(x_i) | = \infty (\forall i) \right\},$$

可见当 $u \in A^c$ 时 u 的 Fourier 级数在每点 x_i 发散. 这就得出:几乎所有 $u \in C(J)$ 的 Fourier 级数在一稠集上发散. □

如果注意到,要构成一个具体的 $u \in C(J)$, 使其 Fourier 级数在某一给定点发散,并非易事(第一个这样的例子由 Fejér 于 1910 年给出),而在理论上我们却敢断言:这种函数不仅存在,而且极多,那么例 2.4.4 所得结论之深刻就更令人惊异了. 观于此,对于一致有界原理的惊人效力,你能不产生深刻印象吗?

§ 2.4.3　Hahn-Banach 定理

现在转而考虑第三个基本定理,与前两个基本定理不同,它是关于有界线性泛函的,而且**不涉及空间的完备性**.

在理论分析中,对于对偶空间 X^* 经常提出的问题是:

(A)**存在问题**:满足给定条件的 $u \in X^*$ 是否存在?

以上问题的解答通常又依赖于另一个与之关联的问题:

(B)**延拓问题**: X 的子空间上的有界线性泛函能否延拓为 X 上的有界线性泛函且满足一定条件?

Hahn-Banach 定理正是对延拓问题给出了很一般的解答.

以下设 X 是给定的赋范空间. 若 X 上的实泛函 $p(x)$ 满足条件:

(i) **正齐次性**: $p(\alpha x) = \alpha p(x) (\alpha > 0, x \in X)$;

(ii) **次可加性**: $p(x + y) \leqslant p(x) + p(y) (x, y \in X)$,

则称 $p(x)$ 为**次线性泛函**. 显然范数就是次线性泛函.

定理 2.4.5(Hahn-Banach 定理)① 设 A 是 X 的子空间.

---

① 文献中对 Hahn-Banach 定理的说法颇不一致. 有时 Hahn-Banach 定理指定理 2.4.5 中的(i). 实际上,将包括定理 2.4.5 在内的一组互有联系的定理统称为 Hahn-Banach 定理,似乎更为恰当.

(i) 若 u 是 A 上的**实**线性泛函[①], p 是 X 上的次线性泛函, $u(x) \leqslant p(x)(\forall x \in A)$, 则 u 可延拓为 X 上的实线性泛函 v, 使得 $v(x) \leqslant p(x)(\forall x \in X)$.

(ii) 若 u 是 A 上的有界线性泛函,则 u 可延拓为 X 上的有界线性泛函 v, 使得

$$\| v \| = \| u \|_A \triangleq \sup_{x \in A, \| x \| \leqslant 1} | u(x) |. \tag{2.4.5}$$

满足式(2.4.5)的 v 称为 u 在 X 上的**保范延拓**.

证 (i)的证明依赖于一定的集论公理,可参考文献[10] §2.10. 下面只证(ii).

首先设 X 是实赋范空间. 令 $p(x) = \| u \|_A \| x \|$, 显然 $p(x)$ 是 X 上的次线性泛函, $u(x) \leqslant p(x)(\forall x \in A)$. 由(i), u 可延拓为 X 上的线性泛函 v, 使得 $v(x) \leqslant p(x)(\forall x \in X)$. 因对任给 $x \in X$ 有

$$v(x) \leqslant \| u \|_A \| x \|,$$

$$-v(x) = v(-x) \leqslant p(-x) = \| u \|_A \| x \|,$$

放得 $| v(x) | \leqslant \| u \|_A \| x \|$, 从而 $v \in X^*$, $\| v \| \leqslant \| u \|_A$. 另一方面显然 $\| v \| \geqslant \| u \|_A$, 故 v 是 u 的保范延拓.

其次设 X 是复赋范空间,则 $\mathrm{Re}\, u$ 是 A 上的实线性泛函,且

$$\mathrm{Re}\, u(x) \leqslant | u(x) | \leqslant \| u \|_A \| x \| \quad (x \in A).$$

将 X 看作**实**赋范空间[②]应用已证结论,得出 X 上的**实**线性泛函 φ, 它是 $\mathrm{Re}\, u$ 的保范延拓. 令

$$v(x) = \varphi(x) - \mathrm{i}\varphi(\mathrm{i}x) \quad (x \in X),$$

则易直接验证:

$$v(\alpha x + \beta y) = \alpha v(x) + \beta v(y), \quad v(\mathrm{i}x) = \mathrm{i}v(x)$$

$(\alpha, \beta \in \mathbf{R}, x, y \in X)$, 由此推出 v 是 X 上的**复**线性泛函. $\forall x \in A$, 有

$$v(x) = \mathrm{Re}\, u(x) - \mathrm{i}\, \mathrm{Re}\, u(\mathrm{i}x)$$

---

① u 是实线性泛函,指其取实值且满足 $u(\alpha x + \beta y) = \alpha u(x) + \beta u(y)(\alpha, \beta \in \mathbf{R})$. 注意,即使 X 是复向量空间,其上的实线性泛函也是有意义的.

② 任何复空间 X 都可看作实向量空间,因 $\forall \alpha \in \mathbf{R}, \forall x \in X$, 必定有 $\alpha x \in X$. 例如,1 维复向量空间 $\mathbf{C}$ 看作实向量空间时,可等同于 2 维向量空间 $\mathbf{R}^2$.

$$= \text{Re}\, u(x) + \text{i}\, \text{Im}\, u(x) = u(x),$$

故 v 是 u 的延拓. 任给 $x \in X$，有

$$|v(x)| = v(x)\text{sgn}\, v(x) = v(x\, \text{sgn}\, v(x))\text{ ①}$$

$$= \varphi(x\, \text{sgn}\, v(x)) \leqslant \|\varphi\| \|x\| \leqslant \|u\|_A \|x\|.$$

因此如同证明第一段一样，有 $v \in X^*$，$\|v\| = \|u\|_A$. □

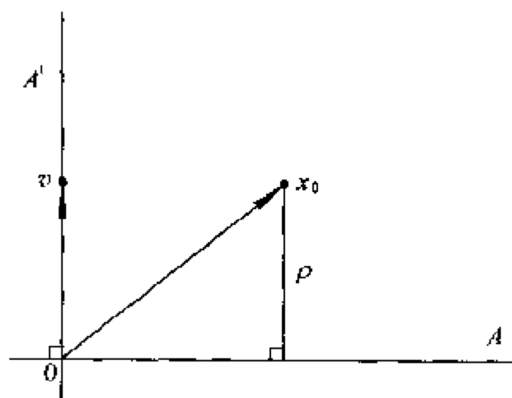


图 2-3

现在应用已证的延拓定理来解决一定线性泛函的存在问题. 问题源于 Hilbert 空间中的以下事实：若 A 是闭子空间，x_0 是 A 外一点，则必有单位向量 $v \in A^\perp$，使得 $\langle v, x_0\rangle = \rho = d(x_0, A)$（见图 2-3）. 以上结论很容易由正交分解定理 1.5.10 推出（请试证）. 现在的问题是：在赋范空间中是否有类似结论？这由以下定理回答.

定理 2.4.6（存在定理）　设 A 是 X 的闭子空间，$x_0 \in X \setminus A$，则存在 $v \in A^-$（记号依式(2.3.14)），使得 $\|v\| = 1, v(x_0) = \rho \triangleq d(x_0, A)$. 特别，取 $A = \{0\}$ 得出：若 $0 \neq x_0 \in X$，则存在 $v \in X^*$，使得 $\|v\| = 1, v(x_0) = \|x_0\|$.

证　令 $B = A + \mathbf{K}x_0$，则 B 是 X 的子空间；因显然 $A \cap \mathbf{K}x_0 = \{0\}$，故 $B = A \oplus \mathbf{K}x_0$. 定义 B 上的泛函 u：

$$u(a + \lambda x_0) = \lambda\rho \quad (a \in A, \lambda \in \mathbf{K}),$$

则 u 显然是 B 上的线性泛函，$u(a) = 0 (\forall a \in A), u(x_0) = \rho$；

$$\|u\|_B = \sup_{0 \neq x \in B} \frac{|u(x)|}{\|x\|} = \sup_{a \in A, 0 \neq \lambda \in \mathbf{K}} \frac{|\lambda| \rho}{\|a + \lambda x_0\|}$$

$$= \sup_{a \in A, 0 \neq \lambda \in \mathbf{K}} \frac{\rho}{\|x_0 + \lambda^{-1} a\|}$$

$$= \frac{\rho}{\inf\limits_{a \in A} \|x_0 - a\|} = \frac{\rho}{d(x_0, A)} = 1.$$

上面用到 $\rho > 0$（何故?）. 设 v 是 u 在 X 上的保范延拓（依定理 2.4.5），则 $\|v\| = \|u\|_B = 1$，

① 参看第 77 页注①.

$$v(x_0) = u(x_0) = \rho, \quad v(a) = u(a) = 0(\forall a \in A).$$

可见 v 正合于定理要求. □

初看起来,定理 2.4.6 的结论与证明似乎皆无惊人之处,但实际上它已为一套系统地使用的方法奠定了基础. 关键之点是 2.4.6 蕴涵了如下重要结论:

推论 2.4.7　$\forall x, y \in X, x \neq y \Leftrightarrow \exists u \in X^*$, 使得 $u(x) \neq u(y)$; 或等价地,

$$x = y \Leftrightarrow \forall u \in X^*, \text{有 } u(x) = u(y) \tag{2.4.6}$$

证　若 $x = y$, 则当然 $u(x) = u(y)(\forall u \in X^*)$. 反之,若 $x \neq y$, 则由定理2.4.6有 $u \in X^*$, 使得

$$u(x - y) = \| x - y \| \neq 0,$$

从而 $u(x) \neq u(y)$. □

推论 2.4.7 的价值在于,它为判定赋范空间中两向量相等提供了一种普遍方法,其形式极为简单:若要证**向量**等式 $x = y$ (这通常是一个无限维问题),则只要证**数量**等式

$$u(x) = u(y) \quad (\forall u \in X^*),$$

后者往往是一个初等问题.

应用以上思想的一个成功且有趣的例证,是无限维向量分析的建立. 所谓**向量分析**就是(实变)向量函数的微积分学,其基本概念的定义形式上与经典微积分学并无区别. 设 $x(\cdot):[a,b] \to X$ 是一向量值函数. 完全照搬经典微积分学中的定义,令

$$x'(t) = \lim_{\tau \to 0} \frac{x(t+\tau) - x(t)}{\tau} \quad (t \in a, b); \tag{2.4.7}$$

$$\int_a^b x(t)\mathrm{d}t = \lim_{\max \Delta t_i \to 0} \sum_{i=1}^{n} x(\tau_i)\Delta t_i, \tag{2.4.8}$$

其中 $a = t_0 < t_1 < \cdots < t_n = b$ 是 $[a,b]$ 的任一分划, $\tau_i \in [t_{i-1}, t_i]$, $\Delta t_i = t_i - t_{i-1}$. 式(2.4.7)与式(2.4.8)的左端分别称为函数 $x(t)$ 的导数与 Riemann 积分. 任给 $u \in X^*$, 容易验证

$$u(x'(t)) = [u(x(t))]', \tag{2.4.9}$$

$$u\left(\int_a^b x(t)\mathrm{d}t\right) = \int_a^b u(x(t))\mathrm{d}t, \tag{2.4.10}$$

只要等式左端有意义. 形式上,以上两式可解释为:

有界线性泛函与微(或积)分运算可交换.

若 X 完备，$x(t)$ 连续，则如微积分学中一样可证明式(2.4.8)中的积分存在.

定理 2.4.8(Newton-Leibniz 公式)　设 X 是一个 Banach 空间，函数 $x(\cdot):[a,b]\to X$ 有连续导数，则成立

$$\int_a^b x'(t)\mathrm{d}t = x(b)-x(a). \tag{2.4.11}$$

证　等式(11)两边有确定意义且都是 X 中的向量，因此，这是一个证明向量等式的问题，正好应用推论 2.4.7. $\forall u\in X^*$，有

$$\begin{aligned} u\left(\int_a^b x'(t)\mathrm{d}t\right) &= \int_a^b u(x'(t))\mathrm{d}t \qquad (\text{用式}(2.4.10)) \\ &= \int_a^b [u(x(t))]'\mathrm{d}t \qquad (\text{用式}(2.4.9)) \\ &= u(x(b)) - u(x(a)) \quad (\text{用 Newton-Leibniz 公式}) \\ &= u(x(b)-x(a)), \end{aligned}$$

于是由推论 2.4.7 知式(2.4.11)为真. □

上面的论证是如此简捷，可能会使你不胜惊讶. 实际上，诸如分部积分公式、变量代换公式等结果都可同样轻而易举地推广到 Banach 空间中. 我们不打算再举这类近乎平凡的例子，只是指出：借助于推论 2.4.7，可将经典微积分学中许多结论，提升为 Banach 空间中的相应结论，因而避开对结论的直接证明(直接证明未必是平凡的). 这种方法应用如此普遍且高度标准化，以至在很多情况下，几乎是例行的标准程序完全可以省去. 这样，几乎没费什么力气，就可以认为已确立了 Banach 空间中的向量分析. 然而你应记住，这一切的基础是 Hahn-Banach 定理！鉴于此，对于 Hahn-Banach 定理的应用价值，你大概已有初步印象了.

§2.4.4　再论对偶空间

在§2.3 中，我们求得了若干赋范空间上的有界线性泛函的通式. 不过，很难指望对一般的赋范空间获得像定理 2.3.2～2.3.6 这样明确的结果. 这就产生一个问题：如果对于 X^* 的结构都不甚了解，还能将其应用于对 X 的研究吗？实际上，在 X^* 的许多应用中，未必需要关于 X^* 的太多信息；特别，未必需要 $u\in X^*$ 的具体表示. 在很多情况下，我们所需要的关于 X^* 的知识，仅限于 Hahn Banach 定理而已. 下面就来说明，应用 Hahn-Banach 定理(当然包括其推论)，能获得关于 X 与 X^* 的哪些一般结论.

定理 2.4.9　设 $\varnothing\neq A\subset X$，则 A 是 X 的基本集 $\Leftrightarrow A^\perp=\{0\}$.

证　若 A 是基本集，则 $\overline{\mathrm{span}A}=X$. 任取 $u\in A^\perp$，直接看出 $u\in(\mathrm{span}A)^\perp$.

$\forall x \in X$, 取 $\{x_n\} \subset \mathrm{span}A$, 使 $x_n \to x$, 则 $u(x) = \lim_n u(x_n) = 0$. 因此 $u = 0$, 故得 $A^\perp = \{0\}$. 反之,若 A 不是基本集,则有 $x_0 \in X \setminus \overline{\mathrm{span}A}$. 由定理 2.4.6,有 $v \in (\overline{\mathrm{span}A})^\perp$, 使得 $v(x_0) \neq 0$, 从而 $v \neq 0$. 必定 $v \in A^\perp$, 因此得 $A^\perp \neq \{0\}$. □

你不妨回忆一下定理 1.5.5,它包含了这样的结论:若 $A = \{e_i\}$ 是 Hilbert 空间 H 中的标准正交系,则 A 为标准正交基 $\Leftrightarrow A^\perp = \{0\}$. 定理 2.4.9 可看作上述结论在赋范空间中的推广.

利用定理 2.4.9 来判定基本集有时十分简便.例如,设 $1 \leqslant p = q/(q-1) < \infty$, $A = \{\chi_{[a,x]} : x \in J\}$, $J = [a,b](a < b)$, 今指明 A 是 $L^p(J)$ 的基本集(参考例 1.3.8(iii)).事实上,若 $v \in A^\perp$, 则 $v \in L^q(J)$ (用定理 2.3.3),且

$$\int_a^x v(y)\mathrm{d}y = 0 \quad (\forall x \in J),$$

对上式求导得 $v(x) = 0$,a.e.. 因此 $A^\perp = \{0\}$, 从而 A 是 $L^p(J)$ 的基本集.

定理 2.4.10 X 可等距嵌入到 X^{**} 中.

证 对任给 $x \in X$, 定义

$$\tilde{x}(u) = u(x) \quad (u \in X^*), \tag{2.4.12}$$

则直接看出 $\tilde{x} \in X^{**}$, 且 $\|\tilde{x}\| \leqslant \|x\|$. 当 $x = 0$ 时显然 $\tilde{x} = 0$. 若 $x \neq 0$, 则由定理 2.4.6 有 $u \in X^*$, 使 $\|u\| = 1, u(x) = \|x\|$, 因此

$$\|x\| = \tilde{x}(u) \leqslant \|\tilde{x}\| \|u\| = \|\tilde{x}\|,$$

故得 $\|x\| = \|\tilde{x}\|$. 这就得到等距映射

$$T : X \to X^{**}, \quad x \to \tilde{x}, \tag{2.4.13}$$

$\tilde{x}$ 定义如式(2.4.12). 由式(2.4.12)直接看出 T 是线性的,因此, T 就是所需要的等距嵌入(参看定义 1.1.6). □

映射式(2.4.13)称为从 X 到 X^{**} 的**正则嵌入**.若等同 x 与 $\tilde{x}$, 则不妨认为 $X \subset X^{**}$. 这意味着,每个 $x \in X$ 有**双重含义**:一方面,它是 X 中的向量;同时,它又可看作 X^* 上的有界线性泛函(即由式(2.4.12)定义的 $\tilde{x}$). 这种观点初看起来似乎可能滋生混乱,但它凸显了 X 与 X^* 之间的对偶关系,特别富有启发性. 关键的思想在于:既然 $x \in X$ 也可看作有界线性泛函,那么,对于有界线性泛函成立的结论亦可用于**向量** x. 例如,我们不必证明就可断言:

(i) 关于 $x \in X$ 有向量范数公式(参照式(2.3.1)),

$$\|x\| = \sup_{u \in X^*, \|u\| \leqslant 1} |u(x)|. \tag{2.4.14}$$

(ii) 任给 $A \subset X$，A 有界 $\Leftrightarrow \forall u \in X^*$，$u(A)$ 在 $\mathbf{K}$ 中有界(参照定理 1.4.3 (iv)，注意此处**不必要求 X 完备**).

最后一个说明给你留下一个遗憾：X 与 X^* 中的命题并非完全互成对应. X 未必是完备的，而 X^* 则必定完备，这已显示出 X 与 X^* 之间的不对称性. 问题在于，虽然 X^* 是 X 的对偶空间，但 X 并非 X^* 的对偶空间，因而不能说 X 与 X^* 互为对偶，除非 $X = X^{**}$. 仅当 $X = X^{**}$ 时，上述的不对称性才得以消除，致使 X 与 X^* 之间的对偶关系达到很完美的程度. 这样的 X 无疑会有更好的性质，因此作以下定义.

定义 2.4.11　若正则嵌入式(2.4.13)为同构，则称 X 为**自反空间**.

有限维赋范空间、Hilbert 空间与 L^p 空间 $(1 < p < \infty)$ 都是自反空间. 自反空间的优势在于，它在很多方面远比一般赋范空间更接近于 Hilbert 空间. 在 §1.6 中，我们提到空间 $L^p(\Omega)$ 较之 $C^m(\Omega)$ 具有更好的结构，其主要理由就在于：$L^p(\Omega)(1 < p < \infty)$ 是自反空间，而 $C^m(\Omega)$ 则不是.

作为 X^* 的对偶空间，X^{**} 必然是完备的；因此，X 在 X^{**} 中的闭包 $\bar{X}$ (它是 X^{**} 的闭子空间)亦是完备的，故是 X 的完备化. 这就顺便得到了定理 1.1.8 的一个极短的证明(惟一性部分易直接证明).

上节中已考虑过对偶算子概念，现在应用本节介绍的基本定理得出某些进一步的结论. 首先完成等式 $\|T^*\| = \|T\|$ (式(2.3.20))的证明，为此只要证明 $\|T\| \leqslant \|T^*\|$ (参看命题2.3.9 的证明). 设 $x \in X$ 使 $\|x\| = 1$，$Tx \neq 0$，则有 $v \in Y^*$，使 $\|v\| = 1$，$v(Tx) = \|Tx\|$ (用定理 2.4.6). 于是

$$\|Tx\| = (T^*v)(x) \leqslant \|T^*\|\,\|v\|\,\|x\| = \|T^*\|,$$

这推出 $\|T\| \leqslant \|T^*\|$ (用式(2.1.7))，如所要证.

若 $T \in L(X, Y)$，则关于线性方程

$$Tx = y \tag{2.4.14$_y$}$$

与

$$T^*v = u \tag{2.4.15$_u$}$$

的可解性问题出现在许多应用领域. 显然，方程 $(2.4.14)_y$ 有解 $\Leftrightarrow y \in R(T)$，$x$ 是方程$(2.4.14)_0$ 的解 $\Leftrightarrow x \in N(T)$. 由此可见，线性方程的可解性问题密切联系于核与值域的性质. 下面是关于核与值域的一个简单结果.

命题 2.4.12　设 $T \in L(X, Y)$，则

$$N(T) = {}^{\perp}R(T^*), \quad N(T^*) = R(T)^{\perp}, \tag{2.4.16}$$

$$^{\perp}N(T^*) = \overline{R(T)}. \tag{2.4.17}$$

关于零化子的记号依式(2.3.14).

证 对于式(2.4.16)仅证前一式,后者是类似的. $\forall x \in X$,有

$$\begin{aligned} Tx = 0 &\Leftrightarrow \langle v, Tx\rangle = 0(\forall v \in Y^*) &\text{(用推论 2.4.7)}\\ &\Leftrightarrow \langle T^* v, x\rangle = 0(\forall v \in Y^*) &\text{(用式(2.3.16))}\\ &\Leftrightarrow x \in {}^{\perp}R(T^*), &\text{(用式(2.3.14))} \end{aligned}$$

这得出 $N(T) = {}^{\perp}R(T^*)$.

其次证式(2.4.17). 利用式(2.3.16)直接看出 $R(T) \subset {}^{\perp}N(T^*)$. 易见 ${}^{\perp}N(T^*)$ 是闭的,故 $\overline{R(T)} \subset {}^{\perp}N(T^*)$. 为证 $\overline{R(T)} \supset {}^{\perp}N(T^*)$,只要对 $y \in Y \setminus \overline{R(T)}$ 证 $y \overline{\in} {}^{\perp}N(T^*)$. 由定理 2.4.6,有 $v \in \overline{R(T)}^{\perp}$ ($\subset R(T)^{\perp} = N(T^*)$,用式(2.4.16)),使得 $v(y) \neq 0$,因而 $y \overline{\in} {}^{\perp}N(T^*)$,如所要证. □

利用方程(2.4.14),(2.4.15)可将式(2.4.16)解释如下:

(i) x 是方程$(2.4.14)_0$ 的解 $\Leftrightarrow$ x 正交于任何 $T^* v(v \in Y^*)$;

(ii) v 是方程$(2.4.15)_0$ 的解 $\Leftrightarrow$ v 正交于任何 $Tx(x \in X)$.

这些事实在积分方程理论中是熟知的.

对等式(2.4.17)亦可作类似的解释. 不过,你可能注意到,(2.4.17)显得有些不和谐,它似乎应改成 ${}^{\perp}N(T^*) = R(T)$,而且还应有一个等式 $N(T)^{\perp} = R(T^*)$ 与之配对. 可惜,这种设想只有在一定附加条件下才能实现(参考定理2.4.13与定理3.3.7).

在线性代数中有以下熟知结论:若 $A \in \mathbf{K}^{m\times n}$,则 $\operatorname{rank} A = m \Leftrightarrow$ 线性方程 $A^{T}x = 0$ 只有零解. 换成算子的说法就是: A 为满射 $\Leftrightarrow N(A^{T}) = \{0\}$. 下面是这一事实的无限维推广.

定理 2.4.13(满射定理) 设 X, Y 是 Banach 空间, $T \in L(X, Y)$. 则 $R(T) = Y \Leftrightarrow T^*$ 有有界逆; $R(T^*) = X^* \Leftrightarrow T$ 有有界逆;若 T 是满射,则 $R(T^*) = N(T)^{\perp}$;若 T^* 是满射,则 $R(T) = {}^{\perp}N(T^*)$.

注 此处 T 有有界逆意指: T 是单射且 $T^{-1}: R(T) \to X$ 有界(参考命题2.1.8).

§2.4.5 分离定理

现在介绍 Hahn Banach 定理的另一组推论,它们具有明显的几何特征,在涉及凸性的问题中有广泛应用.

设 X 是实赋范空间. 任给 $a, b \in X$,称点集

$$[a, b] \triangleq \{ta + (1-t)b : 0 \leqslant t \leqslant 1\} \tag{2.4.18}$$

为 X 中以 a, b 为端点的**线段**. 若 $A \subset X$, $\forall a, b \in A$,有 $[a, b] \subset A$,则称 A 为

凸集. 常见的凸集有：

(i) 子空间；子空间的平移象(即形如 $a + A$ 的集，A 是子空间).

(ii) 半空间：设 $0 \neq u \in X^*$，$\beta \in \mathbf{R}$，则形如 $X(u \geqslant \beta)$，$X(u \leqslant \beta)$，$X(u > \beta)$ 与 $X(u < \beta)$ 的集都称为**半空间**.

(iii) 开球与闭球.

若 $a_i \in X$，$\lambda_i \geqslant 0 (1 \leqslant i \leqslant n)$，$\sum \lambda_i = 1$，则称 $\sum \lambda_i a_i$ 为 $\{a_i\}$ 的**凸组合**. 例如，由(18)定义的线段 $[a, b]$ 中每点都是 $\{a, b\}$ 的凸组合. 用归纳法可以证明：$A \subset X$ 是凸集 $\Leftrightarrow A$ 中任何有限个点的凸组合均属于 A.

设 $A, B \subset X$，$0 \neq u \in X^*$，$r \in \mathbf{R}$. 若

$$u(A) \leqslant r \leqslant u(B), \tag{2.4.19}$$

即 $u(a) \leqslant r \leqslant u(b) (\forall a \in A, \forall b \in B)$，则集 A 与 B 分别位于闭的半空间

$$X(u \leqslant r) \text{ 与 } X(u \geqslant r)$$

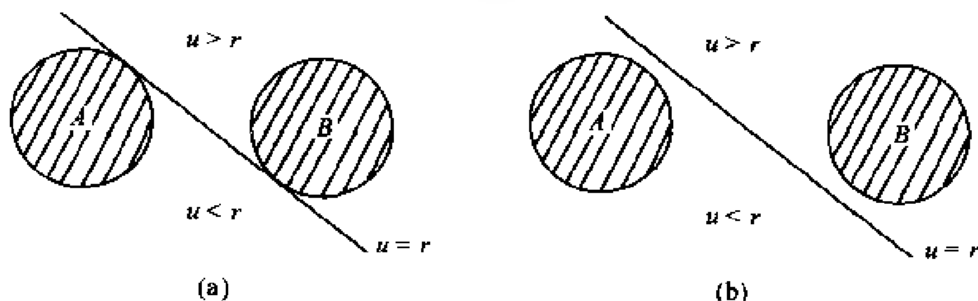


图 2-4

中，此时说超平面 $X(u = r)$ **分离**集 A 与 B (图 2-4(a)). 若

$$u(A) < r < u(B), \tag{2.4.20}$$

即 $u(a) < r < u(b) (\forall a \in A, \forall b \in B)$，则集 A 与 B 分别位于开的半空间

$$X(u < r) \text{ 与 } X(u > r)$$

中，此时说超平面 $X(u = r)$ **严格分离** A 与 B (图 2-4(b)). 直观上很明显，即使 $A \cap B = \varnothing$，A 与 B 亦未必可用超平面分离. 可分离性强烈地依赖于 A 与 B 的凸性. 下面是一个基本的结果.

定理 2.4.14(分离定理)　设 $A, B \subset X$ 是非空凸集.

(i) 若 $A^\circ \neq \varnothing$，$A^\circ \cap B = \varnothing$，则存在 $u \in X^*$，$r \in \mathbf{R}$，使得

$$u(A^\circ) < r \leqslant u(B). \tag{2.4.21}$$

(ii) 若 A 为紧集, B 为闭集, $A\cap B=\varnothing$, 则存在 $u\in X^*$, $r\in\mathbf{R}$, 使得不等式(2.4.20)成立.

证 (i) 证明的关键是应用 Hahn-Banach 定理,因包含一些技术性细节而略嫌过长,故从略(参考文献[10]§2.10).

(ii) 证明的基本想法是用两个凸开集将 A 与 B 分离,然后应用(i).令

$$U=\{x\in X: d(x,A)<\varepsilon\},\quad V=\{x\in X: d(x,B)<\varepsilon\}.$$

显然 $A\subset U$, $B\subset V$, U 与 V 是开集(参考第一章习题16).稍细的验证知 U 与 V 亦是凸集.选取 ε 充分小,使 $U\cap V=\varnothing$. 这种 ε 必存在,否则,必有 $\{x_n\}\subset X$, 使 $d(x_n,A)\to 0$, $d(x_n,B)\to 0(n\to\infty)$. 而这又推出存在 $\{a_n\}\subset A$, $\{b_n\}\subset B$, 使 $\|a_n-b_n\|\to 0(n\to\infty)$. 因 A 为紧集, B 为闭集,不妨设 $a_n\to a\in A$, $b_n\to a\in B(n\to\infty)$, 从而 $a\in A\cap B$, 而这矛盾于 $A\cap B=\varnothing$! 因此 U 与 V 是互不相交的开凸集.由(i),有 $u\in X^*$, $r\in\mathbf{R}$, 使得 $u(U)<r\leqslant u(V)$. 必 $r<u(V)$ (从而不等式(2.4.20)成立).否则,有 $x\in V$ 使 $u(x)=r$. 必有 $z\in X$ 使 $u(z)<0$ (何故?).因 V 为开集,故当 $\delta>0$ 充分小时 $x+\delta z\in V$, 而

$$u(x+\delta z)=r+\delta u(z)<r,$$

这矛盾于 $r\leqslant u(V)$. □

通常认为,分离定理是 Hahn-Banach 定理的几何形式,它的应用如同 Hahn-Banach 定理本身一样广泛.在凸分析、最优化理论等数学分支中,分离定理的应用尤其常见.然而令人为难的是,眼下却难以举出特别有趣的例子,除非先准备一大堆概念与预备知识,这显然不是你乐意看到的.不过,以下的例子还不算复杂且颇有意义.

例 2.4.15(对偶锥) 若 $K\subset X$ 满足 $tK\subset K(\forall t>0)$, 则称 K 为**锥**,称

$$K^*\triangleq\{u\in X^*: u(K)\geqslant 0\}$$

为 K 的**对偶锥**,令 $K^{**}=(K^*)^*$. 若 X 是实自反空间, $K\subset X$ 是锥且为闭凸集,则 $K=K^{**}$.

证 不难看出 $K\subset K^{**}$, 关键是证 $K\supset K^{**}$. 由 X 自反得出 $K^{**}\subset X$. 设 $x\in X\setminus K$, 只要证 $x\overline{\in} K^{**}$, 即有 $u\in K^*$ 使 $u(x)<0$. 分别以 $\{x\}$ 与 K 代 A 与 B 应用定理 2.4.14(ii),得出 $u\in X^*$ 与 $r\in\mathbf{R}$, 使得 $u(x)<r<u(K)$. 任取 $y\in K$ 与 $t>0$, 有 $ty\in K$. 因 K 是闭集,令 $t\to 0$ 得 $0\in K$. 于是 $r<u(0)=0$. 从而 $u(x)<0$. 余下只要证 $u\in K^*$. 用反证法:若 $u\overline{\in} K^*$, 则有 $z\in K$, 使得 $u(z)<0$, 于是

$$-\infty=\lim_{t\to\infty}tu(z)=\lim_{t\to\infty}u(tz)\geqslant r,$$

这显然不可能.

§2.5 弱 收 敛

在同一函数类(例如连续函数类)中,可以考虑多种不同的收敛性(例如一致收敛、平均收敛与点态收敛等),这是你在§1.2中已看到了的.现在要指出,即使在一个抽象的赋范空间 X 中,亦可考虑范数收敛之外的其他收敛.范数收敛固然有良好的性质,但往往要求过强.例如,空间 $C(J)$ 中的范数收敛就是一致收敛.我们知道,即使对于微积分学中的一些问题,如积分号下取极限,一致收敛也不是必要的.因此,考虑弱于范数收敛的收敛性有其现实意义.

§2.5.1 弱收敛及其刻画

下面平行地考虑 X 与 X^* 中的弱收敛.

定义 2.5.1　设 $\{x_n\}\subset X$ 与 $\{u_n\}\subset X^*$ 是两个序列.

(i) 若 $x\in X$, $\forall u\in X^*$, 有 $u(x_n)\to u(x)(n\to\infty)$, 则说 $\{x_n\}$ **弱收敛**于 x, 记作 $x_n\rightharpoonup x(n\to\infty)$.

(ii) 若 $u\in X$, $\forall x\in X$, 有 $u_n(x)\to u(x)(n\to\infty)$, 则说 $\{u_n\}$ **弱*收敛**于 u, 记作 $u_n\overset{*}{\rightharpoonup} u(n\to\infty)$.

显然,弱*收敛就是在 X 上的点态收敛.若 X 是自反空间,则 X^* 中的弱*收敛就是弱收敛.在一般情况下,弱*收敛可能弱于弱收敛.

直接看出, $x_n\to x\Rightarrow x_n\rightharpoonup x$; $u_n\to u\Rightarrow u_n\overset{*}{\rightharpoonup} u$, 即范数收敛强于弱收敛与弱*收敛.因此,也将范数收敛称为**强收敛**.若 $\dim X<\infty$, 例如设 $X=\mathbf{K}^n$, 则在 X 中

$$
\begin{aligned}
x^{(k)}\rightharpoonup x &\Rightarrow \langle x^{(k)},e_i\rangle\to\langle x,e_i\rangle\\
&\Rightarrow x_i^{(k)}\to x_i\\
&\Rightarrow x^{(k)}\to x\quad(1\leqslant i\leqslant n)\quad(k\to\infty),
\end{aligned}
$$

其中 $\{e_i\}$ 是 $\mathbf{K}^n$ 的标准基.可见 $x^{(k)}\rightharpoonup x\Leftrightarrow x^{(k)}\to x(k\to\infty)$. 因此,在有限维赋范空间中弱收敛与强收敛没有区别.

下面对弱收敛给出**形式上更弱的条件**,以使实际判定弱收敛时更加容易.在以下定理中,基本集概念是关键的(参考定义1.3.7(iii)).

定理 2.5.2　设 B 与 G 分别为 X 与 X^* 的基本集.

(i) 设 $x,x_n\in X(n=1,2,\cdots)$. 则 $x_n\rightharpoonup x$ 的充要条件是

$$\sup_n\|x_n\|<\infty \text{ 且 } u(x_n)\to u(x)(n\to\infty,\forall u\in G). \tag{2.5.1}$$

(ii) 设 $u, u_n \in X^*$ $(n = 1,2,\cdots)$. 则 $u_n \xrightarrow{*} u$ 的充分条件是:

$$\sup_n \| u_n \| < \infty \text{ 且 } u_n(x) \to u(x)(n \to \infty, \forall x \in B). \qquad (2.5.2)$$

当 X 完备时,式(2.5.2)是 $u_n \xrightarrow{*} u$ 的充要条件.

证 (i) 若 $x_n \rightharpoonup x$, 则 $\forall u \in X^*$, $\{u(x_n)\}$ 有界,于是由一致有界原理推出 $\{x_n\}$ (看作 X^{**} 中的序列)有界.因此条件(1)满足.

其次,设条件(2.5.1)式满足,令 $\beta = \sup_n \| x_n \|$. 任给 $u \in X^*$, 要证 $u(x_n) \to u(x)(n \to \infty)$. $\forall \varepsilon > 0$, 取 $v \in \text{span } G$, 使 $\| u - v \| < \varepsilon$. 条件(2.5.1)式显然推出 $v(x_n) \to v(x)(n \to \infty)$, 因此有 $N > 0$, 使得

$$| v(x_n) - v(x) | < \varepsilon \quad (\forall n \geqslant N).$$

于是,当 $n \geqslant N$ 时,有

$$\begin{aligned} | u(x_n) - u(x) | &\leqslant | u(x_n) - v(x_n) | + | v(x_n) - v(x) | \\ &\quad + | v(x) - u(x) | \\ &< \| u - v \| (\| x_n \| + \| x \|) + \varepsilon \\ &< \varepsilon(\beta + \| x \| + 1), \end{aligned}$$

这表明 $u(x_n) \to u(x)(n \to \infty)$, 如所要证.

(ii) 的证明是类似的,仅有的区别是,在证 $u_n \xrightarrow{*} u \Rightarrow \sup_n \| u_n \| < \infty$ 时,为用一致有界原理 2.4.3(iv),要求 X 完备. □

简单地说,定理 2.5.2 表明,为判定有界序列弱收敛或弱*收敛,只要判定它在某个基本集上点态收敛就够了.而基本集可以取得尽可能地"小"(例如是一可数集),因而弱收敛(或弱*收敛)条件有可能表达得很简单.特别,对于常用的空间 L^p 与 $C(J)$ (依§1.2),可以由定理 2.5.2 推出,其中的弱收敛具有人们熟知的形态.

例 2.5.3 (i) 空间 $l^p(1 < p < \infty)$ 中的弱收敛.设 $q = p/(p-1)$, $\{e_i\}$ 是 $l^q(= (l^p)^*)$ 的标准基,因而也是 l^q 的基本集(例 1.3.8(i)). 注意到对任给 $x = (x_i) \in l^p$, 有 $\langle e_i, x \rangle = x_i$ (记号依式(2.3.13)),由定理 2.5.2 得出结论:在 l^p 中 $x^{(n)} \to x$ 的充要条件是

$$\sup_n \| x_n \|_p < \infty \text{ 且 } x_i^{(n)} \to x_i \quad (n \to \infty, \forall i \in \mathbf{N}). \qquad (2.5.3)$$

换言之,**l^p 中有界序列的弱收敛亦即按坐标收敛.**

(ii) 空间 $L^p(J)$ 中的弱收敛($1 < p < \infty, J = [a,b], a < b$). 设 $q = p/(p-1)$, 取 $L^q(J)(= L^p(J)^*)$ 的基本集 $\{\chi_{[a,x]} : x \in J\}$ (参考例 1.3.8(iii)),则由

定理 2.5.2 得出结论：在 $L^p(J)$ 中 $u_n \rightharpoonup u$ 的充要条件是

$$\sup_n \| u_n \|_p < \infty \text{ 且} \int_a^x u_n(y)\mathrm{d}y \to \int_a^x u(y)\mathrm{d}y \quad (n \to \infty, x \in J). \tag{2.5.4}$$

(iii) Hilbert 空间中的弱收敛. 设 H 是一Hilbert空间，$\{e_i : i \in \mathbf{N}\}$ 是其标准正交基，记 $\hat{x}_i = \langle x, e_i \rangle$. 则由定理 2.5.2 推出：在 H 中 $x^{(n)} \rightharpoonup x$ 的充要条件是

$$\sup_n \| x^{(n)} \| < \infty \text{ 且 } \hat{x}_i^{(n)} \to \hat{x}_i \quad (n \to \infty, \forall i \in \mathbf{N}). \tag{2.5.5}$$

换言之，**H 中有界序列的弱收敛即按正交坐标收敛**. 由此特别推出，序列 $\{e_i\}$ 弱收敛于零，但它显然不强收敛于零. 将此结论用于三角函数系(式(1.5.18))得出，在 $L^2[-\pi,\pi]$ 中 $\sin nx \rightharpoonup 0(n \to \infty)$，但可指出在 $(0,\pi)$ 内 $\sin nx$ 处处不收敛于 0.

(iv) 空间 $C(J)$ 中的弱收敛($J = [a,b], a < b$). 今指出，在 $C(J)$ 中 $u_n \rightharpoonup u$ 的充要条件是

$$\sup_n \| u_n \|_0 < \infty \text{ 且 } u_n(x) \to u(x) \quad (n \to \infty, \forall x \in J). \tag{2.5.6}$$

换言之，**$C(J)$ 中有界序列的弱收敛即点态收敛**. 事实上，若 $u_n \rightharpoonup u$，$\forall x \in J$，定义 $f_x(u) = u(x)(u \in C(J))$，则 $f_x \in C(J)^*$，因而 $f_x(u_n) \to f_x(u)$，即 $u_n(x) \to u(x)(n \to \infty)$. 反之，若条件(2.5.6)式满足，则 $\forall v \in \mathrm{BV}_0(J)(= C(J)^*$，依定理2.3.4)，有

$$\lim_n \int_a^b u_n(x)\mathrm{d}v(x) = \int_a^b u(x)\mathrm{d}v(x) \quad \text{(用控制收敛定理)},$$

故得 $u_n \rightharpoonup u$.

现在再回到一般的弱收敛概念. 如同强收敛一样，弱收敛的极限是惟一的(请试证)，且具有通常极限的运算性质(如 $x_n \rightharpoonup x, y_n \rightharpoonup y \Rightarrow x_n + y_n \rightharpoonup x + y$)，这些都容易直接用定义 2.5.1 验证. 但关于弱收敛的深入讨论并不简单，且显示出与强收敛有重大差别，这些都不能在此详述. 下面只不加证明地叙述一个最重要的结果，它不仅应用广泛，而且显示出弱收敛与强收敛的深刻差别.

定理 2.5.4 自反 Banach 空间中任何有界序列包含弱收敛子列.

显然，定理中的"弱收敛子列"绝不能改为"强收敛子列"，除非空间是有限维的(参看定理 1.4.8). 定理 2.5.4 的意义在于，它表明自反空间中的有界集在弱收敛意义上具有某种"紧性". 这样，就可将一些基于紧性的结果推广到弱收敛的情况. 你很快就要看到这样的例子(如见定理 2.5.6).

§2.5.2 某些应用

弱收敛概念的应用可分为两类,其一是用于研究收敛性,这通常涉及某个特定的有界线性泛函序列,并需应用定理 2.5.2;其二是用于某些需要弱收敛条件的论证,这类应用并不涉及特定的弱收敛序列.对于以上两类应用,下面各举一例加以说明.

例 2.5.5(求积公式的收敛性) 给定区间 $J=[a,b]$ 的分点 $\{x_k^n\}$:

$$a \leqslant x_0^n < x_1^n < \cdots < x_n^n \leqslant b, \quad n=1,2,\cdots$$

与实系数 $A_k^n(0 \leqslant k \leqslant n, n=1,2,\cdots)$,称

$$\int_a^b u(x)\mathrm{d}x \approx \sum_{k=0}^n A_k^n u(x_k^n) \quad (u \in C(J)) \tag{2.5.7}$$

为求积公式若

$$\int_a^b u(x)\mathrm{d}x = \lim_n \sum_{k=0}^n A_k^n u(x_k^n) \quad (\forall\, u \in C(J)), \tag{2.5.8}$$

则称求积公式(2.5.7)收敛.今要求出式(2.5.7)收敛的条件.为此,分别以 $f(u)$ 与 $f_n(u)$ 记式(2.5.7)之两端,则 $f, f_n \in C(J)^*$ (参考式(2.3.10)与例 2.3.5(ii)),而式(2.5.8)即 $f_n(u) \to f(u)(n \to \infty, u \in C(J))$,这正意味着 $f_n \xrightarrow{*} f$.为应用定理 2.5.2,取 $C(J)$ 的基本集 $B=\{x^k : k \in \mathbf{Z}_+\}$,于是

$$f_n \xrightarrow{*} f \Leftrightarrow \sup_n \| f_n \| < \infty \text{ 且 } f_n(x^k) \to f(x^k) \quad (n \to \infty, k \geqslant 0).$$

现在假定公式(2.5.7)有 n 阶代数精度,这意味着当 u 是次数 $\leqslant n$ 的多项式时式(2.5.7)成为严格等式,则 $f_n(x^k)=f(x^k)(n \geqslant k)$,于是由定理 2.5.2(ii)有

$$f_n \xrightarrow{*} f \Leftrightarrow \sup_n \| f_n \| < \infty.$$

用类似于例 2.3.5(ii)的方法可以得出

$$| f_n | = \sum_{k=0}^n | A_k^n |.$$

因此有结论:若求积公式(2.5.7)有 n 阶代数精度,则它收敛的充要条件是

$$\sup_n \sum_{k=0}^n | A_k^n | < \infty. \tag{2.5.9}$$

特别,若 $A_k^n \geqslant 0$,则取 $u \equiv 1$ 得

$$\sum_{k=0}^{n} |A_k^n| = f_n(u) = f(u) = b - a,$$

可见此时条件(2.5.9)式平凡地满足,因而求积公式必收敛.

由例 2.5.5 直接得出结论:数值积分公式中的梯形公式、Simpson 公式、Cotes 公式等都是收敛的.你从这些事实看出,用泛函分析方法所达到的结论具有多大的普遍价值.

现在考虑应用弱收敛的第二个例子.

定理 2.5.6 设 X 是自反实 Banach 空间, $D \subset X$ 是有界闭凸集, $f: D \to \mathbf{R}$ 在以下意义上连续:

$$x_n \rightharpoonup x\,(x_n, x \in D) \Rightarrow f(x) \leqslant \varliminf_n f(x_n). \tag{2.5.10}$$

则 $f(x)$ 在 D 上取得最小值.

证 证明基本上是定理 1.4.4(ii)之证的一个改制.令 $\alpha = \inf\limits_{x \in D} f(x)$.取序列 $\{x_n\} \subset D$, 使 $f(x_n) \to \alpha\,(n \to \infty)$.由定理 2.5.4,不妨设 $x_n \rightharpoonup x$.必定 $x \in D$, 否则由分离定理 2.4.14 有 $u \in X^*$, $r \in \mathbf{R}$, 使得

$$u(x) < r < u(D).$$

另一方面,由 $x_n \rightharpoonup x$ 推出 $u(x_n) \to u(x)$, 得出矛盾.于是由条件(2.5.10)式得

$$f(x) \leqslant \varliminf_n f(x_n) = \alpha \leqslant f(x),$$

这推出 $f(x) = \alpha$, 从而 $f(x)$ 是 f 在 D 上的最小值. □

定理 2.5.6 为说明弱收敛的作用提供了一个典型例子.在泛函分析中,要考虑包括弱收敛在内的多种收敛性,你可能会为此而感到困惑.或许以为,较强的收敛更有价值,而弱收敛即使有用也相形见绌.其实这是一种误解.从逻辑上说,如果收敛性出现在命题结论中,那么它自然越强越好;相反,如果收敛性出现于命题条件中,那它就越弱越好了.在很多情况下,一种较弱的收敛可能已包含问题所需的信息,而验证收敛条件又容易得多.鉴于此,你就能够理解,有多种收敛性(包括较弱的收敛性)可供选择,其实是一件很有益的事.

§2.5.3 算子序列的收敛性

现在将弱收敛思想用于算子序列.设 $T, T_n \in L(X, Y)\,(n = 1, 2, \cdots)$.对 $\{T_n\}$ 可考虑以下三种收敛性:

一致收敛: $\|T_n - T\| \to 0$, 即依算子范数收敛;

强收敛: $T_n x \to Tx\,(\forall x \in X)$, 即点态收敛;

弱收敛: $T_n x \rightharpoonup Tx\,(\forall x \in X)$.

“一致收敛”一词，不免使人联想到通常意义下的一致收敛. 二者确有联系：序列 $\{T_n\} \subset L(X,Y)$ 一致收敛于 $T \in L(X,Y)$ 的充要条件是，在任何有界集 $B \subset X$ 上 $\|T_n x - Tx\| \rightrightarrows 0 (n \to \infty)$. 若 $Y = \mathbf{K}$，则此处所说的强收敛与弱收敛一致，二者均重合于定义 2.5.1 意义下的弱*收敛. 在一般情况下，显然一致收敛 $\Rightarrow$ 强收敛 $\Rightarrow$ 弱收敛，三者可能互不相同.

例 2.5.7 (i) 设 $T_n \in L(l^2)(n = 1,2,\cdots)$ 定义为

$$T_n x = \sum_{i=n}^{\infty} x_i e_i \quad (x = (x_i) \in l^2),$$

其中 $\{e_i\}$ 是 l^2 的标准基. 显然 $T_n x \to 0 (n \to \infty, \forall x \in l^2)$，即 $\{T_n\}$ 强收敛于零算子. 另一方面，由 $\|T_n x\|_2 \leqslant \|x\|_2$ 及 $T_n e_n = e_n$ 得出 $\|T_n\| = 1$，可见 $\{T_n\}$ 并不一致收敛于 $T = 0$.

(ii) 设 $T_n \in L(l^2)(n = 1,2,\cdots)$ 定义为

$$T_n x = x_1 e_n \quad (x = (x_i) \in l^2),$$

则显然有 $T_n x \xrightarrow{w} 0$（参看例 2.5.3(i)），因而序列 $\{T_n\}$ 弱收敛于 $T = 0$. 但 $T_n e_1 = e_n \not\to 0 (n \to \infty)$，因此 $\{T_n\}$ 不强收敛于 $T = 0$.

以下结果可看作定理 2.5.2(ii)对算子序列的推广.

定理 2.5.8 设 X 与 Y 是 Banach 空间，$\{T_n\} \subset L(X,Y)$，则 $\{T_n\}$ 强收敛的充要条件是：

(i) $\{T_n\}$ 一致有界，即 $\sup_n \|T_n\| < \infty$；

(ii) 存在 X 的基本集 B，使得 $\forall x \in B$，$\{T_n x\}$ 收敛.

证 必要性直接从一致有界原理得出，下面证充分性.

设条件(i)，(ii)满足. 从条件(ii)推出，$\forall x \in \operatorname{span} B$，$\{T_n x\}$ 收敛. $\forall x \in X$，$\forall \varepsilon > 0$，取 $y \in \operatorname{span} B$，使 $\|x - y\| < \varepsilon$. 由

$$\begin{aligned}\|T_m x - T_n x\| &\leqslant \|T_m x - T_m y\| + \|T_m y - T_n y\| \\ &\quad + \|T_n y - T_n x\| \\ &< \varepsilon(\|T_m\| + \|T_n\|) + \|T_m y - T_n y\|\end{aligned}$$

及条件(i)推出 $\{T_n x\}$ 是 Cauchy 列；因 Y 完备，故 $\{T_n x\}$ 收敛，以 Tx 记其极限，则

$$Tx = \lim_n T_n x \quad (x \in X). \tag{2.5.11}$$

由式(2.5.11)定义一个算子 $T: X \to Y$，它显然是线性的. 由

$$\| Tx \| = \lim_n \| T_n x \| \leqslant \varliminf_n \| T_n \| \| x \|$$

推出 $T \in L(X,Y)$ 且

$$\| T \| \leqslant \varliminf_n \| T_n \| . \tag{2.5.12}$$

式(2.5.11)表明 $\{T_n\}$ 强收敛于 T. □

算子序列强收敛常常足以推出一些很强的结论,又不像一致收敛那样难以验证,因而被广泛使用.下面仅从一个方面的应用加以说明.

若 $\{T_n\} \subset L(X)$ 强收敛于单位算子 I,令 $x_n = T_n x (x \in X)$,则 $x_n \to x(n \to \infty)$. 这样,$\{T_n\}$ 就提供了一种构造逼近序列的一般方法.例如,取 $X = C(J), J = [a,b](a < b)$,$P$ 是 J 上的多项式之全体,$\{T_n\} \subset L(X)$ 强收敛于单位算子,$R(T_n) \subset P(n = 1,2,\cdots)$,$\forall u \in C(J)$,令 $u_n = T_n u$,则 $\{u_n\}$ 是 J 上一致逼近 u 的多项式序列.而由定理2.5.7,为验证 $\{T_n\}$ 强收敛于单位算子,只需验证$\sup\limits_n \| T_n \| < \infty$ 且 $T_n x^k \rightrightarrows x^k (n \to \infty, \forall k \in \mathbf{Z}_+)$ 就够了.这就得到一种构造一致逼近多项式的一般方法.非常有趣的是,这一方法为一些著名的经典结论提供了新的证法.下面就是一个颇有说服力的例子.

例 2.5.9(Bernstein 多项式) 任给 $u \in C[0,1]$,令

$$u_n(x) = \sum_{k=0}^{n} \binom{n}{k} u\left(\frac{k}{n}\right) x^k (1-x)^{n-k} \triangleq B_n u(x), \tag{2.5.13}$$

则 $u_n(x)$ 是次数 $\leqslant n$ 的多项式,称为 **Bernstein 多项式**.因

$$| u_n(x) | \leqslant \| u \|_0 \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{(n-k)} = \| u \|_0 ,$$

故由式(2.5.13)定义一个 $C[0,1]$ 上的有界线性算子 B_n,它满足 $\| B_n \| \leqslant 1$. 今证 $\{B_n\}$ 强收敛于单位算子.由定理2.5.8,只要对任给 $m \in \mathbf{Z}_+$,证 $B_n x^m \rightrightarrows x^m (n \to \infty)$. 为此对 m 用归纳法.显然 $B_n x^0 \equiv 1$. 设 $m \geqslant 0, B_n x^k \rightrightarrows x^k (n \to \infty, 0 \leqslant k \leqslant m)$,则

$$\begin{aligned} B_n x^{m+1} &= \sum_{k=0}^{n} \binom{n}{k} \left(\frac{k}{n}\right)^{m+1} x^k (1-x)^{n-k} \\ &= \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-k)!} \left(\frac{k}{n}\right)^m x^k (1-x)^{n-k} \\ &= x \left(\frac{n-1}{n}\right)^m \sum_{k=0}^{n-1} \binom{n-1}{k} \left(\frac{k+1}{n-1}\right)^m x^k (1-x)^{n-1-k} \end{aligned}$$

$$= x\left(\frac{n-1}{n}\right)^m \sum_{j=0}^{m} \binom{m}{j} \left(\frac{1}{n-1}\right)^j \sum_{k=0}^{n-1} \binom{n-1}{k} \left(\frac{k}{n-1}\right)^{m-j} x^k (1-x)^{n-1-k}$$

$$\rightrightarrows x \cdot x^m = x^{m+1}, \quad (n \to \infty)$$

如所要证.

在以上的例子中，$\{B_n\}$ 并不一致收敛于单位算子.事实上，若 $\| B_n - I \| \to 0$，则由后面将建立的一个结果(引理 3.1.1)，当 n 充分大时 B_n 是可逆的.但无论 n 取多大，B_n 都不是单射：显然有不同的 $u, v \in C[0,1]$，使得 $u(k/n) = v(k/n)(0 \leqslant k \leqslant n)$，从而 $B_n u = B_n v$. 然而，$\{B_n\}$ 不一致收敛于单位算子这一点，对于 Bernstein 多项式的一致逼近性质并无任何影响.关键在于，我们所需要的是，对**给定的** $u \in C[0,1]$，有 $B_n u(x) \rightrightarrows u(x)$；至于这种收敛是否也对于满足某一不等式 $\| u \|_0 \leqslant \text{const}$ 的 u 一致，则完全不是问题所要求的，大概也不是任何逼近方法所能做到的.

例 2.5.9 中所用的思想还可以大大发挥与引伸，并因此而发展成解决逼近问题的一套系统方法.此处当然不能详细介绍，但从下面的简单例子你能获得某些进一步的印象.

例 2.5.10(光滑化算子) 取定一个函数 $\varphi \in C^m(\mathbf{R})(0 \leqslant m \leqslant \infty)$，使之满足条件：

$$0 \not\equiv \varphi(x) \geqslant 0, \quad \text{supp}\ \varphi \subset [-1,1]. \tag{2.5.14}$$

令 $\varphi_n(x) = c_n \varphi(nx)$，常数 c_n 决定于条件

$$\int_{-\infty}^{\infty} \varphi_n(x) \mathrm{d}x = 1 \qquad (n \geqslant 1). \tag{2.5.15}$$

显然 $\varphi_n \in C^m(\mathbf{R})$ 且 $\text{supp}\ \varphi_n \subset [-1/n, 1/n]$. 利用 φ_n 定义一族卷积算子(参考§2.2.2)：

$$T_n u = \varphi_n * u \quad (u \in L^1(\mathbf{R})), \tag{2.5.16}$$

则可验证以下结论：

(i) $T_n L^1(\mathbf{R}) \subset C^m(\mathbf{R})$. 实际上，对任给 $u \in L^1(\mathbf{R})$，有

$$\frac{\mathrm{d}^k}{\mathrm{d}x^k} T_n u(x) = \int_{-\infty}^{\infty} \varphi_n^{(k)}(x-y) u(y) \mathrm{d}y \quad (0 \leqslant k < m+1).$$

积分号下求导的合理性依据控制收敛定理.

(ii) $T_n \in L(L^1(R))$，且 $\| T_n \| \leqslant \| \varphi_n \|_1 = 1$ (用命题 2.2.5).

(iii) $\{T_n\}$ 强收敛于单位算子.因 $L^1(\mathbf{R})$ 有基本集(参考例 1.3.8(iv))

$$B = \{\chi_\delta : \delta \subset \mathbf{R} \text{ 是有限区间}\},$$

由定理 2.5.8,只要对任给 $\delta = [a,b] \subset \mathbf{R}$, 验证 $T_n\chi_\delta \xrightarrow{L^1} \chi_\delta$. 这由以下推演得出:

$$\begin{aligned}
\| T_n\chi_\delta - \chi_\delta \|_1 &= \int_{-\infty}^{\infty} \left| \int_{-\infty}^{\infty} \chi_\delta(x-y)\varphi_n(y)\mathrm{d}y - \int_{-\infty}^{\infty} \chi_\delta(x)\varphi_n(y)\mathrm{d}y \right| \mathrm{d}x \\
&\leqslant \int_{-\infty}^{\infty} \mathrm{d}x \int_{-\infty}^{\infty} |\chi_\delta(x-y) - \chi_\delta(x)| \varphi_n(y)\mathrm{d}y \\
&= \int_{-\infty}^{\infty} \varphi_n(y)\mathrm{d}y \int_{-\infty}^{\infty} |\chi_\delta(x-y) - \chi_\delta(x)| \mathrm{d}x \\
&= 2\int_{-\infty}^{\infty} |y| \varphi_n(y)\mathrm{d}y = 2\int_{-1/n}^{1/n} |y| \varphi_n(y)\mathrm{d}y \\
&\leqslant \frac{2}{n}\int_{-1/n}^{1/n} \varphi_n(y)\mathrm{d}y = \frac{2}{n} \to 0 (n \to \infty).
\end{aligned}$$

这样,我们就有了一个一般的方法,对每个 $u \in L^1(\mathbf{R})$, 构造出 C^m 函数序列 $\{u_n\}(u_n = T_n(u))$, 使得 $\{u_n\}$ 平均逼近 u. 通过以 u_n 近似代替 u, 就将 u **光滑化**了,因而称 T_n 为**光滑化算子**.

然而余下一个问题:满足条件(2.5.14)式的 C^m 函数是否存在? 若仅要求连续(即取 $m = 0$),则函数

$$\varphi(x) = \max\{1 - |x|, 0\}$$

就是. 稍复杂点,若令

$$g(x) = \begin{cases} 0, & x \leqslant 0; \\ \mathrm{e}^{-1/x}, & x > 0, \end{cases}$$

$$\varphi(x) = \frac{g(1 - |x|)}{g(1 - |x|) + g(|x| - 1/2)},$$

则可验证 $g \in C^\infty(\mathbf{R})$, 从而 $\varphi \in C^\infty(\mathbf{R})$; φ 显然满足条件(2.5.14)式.

习 题

1. 设 $T: X \to Y$ 为线性算子,证明 $\| T \| \leqslant \beta \Leftrightarrow TB_1(0) \subset \bar{B}_\beta(0)$.
2. 设 $a = (a_i) \in l^\infty$, $Tx = (a_i x_i)$, 说明 $T \in L(l^1)$ 且 $T \in L(l^\infty)$, 求 $\| T \|$.
3. 设 $a = (a_i)$, $Tx = (a_i x_i)$. 证明 $T \in L(l^2) \Leftrightarrow a \in l^\infty$; 当 $a \in l^\infty$ 时求 $\| T \|$.
4. 设 $T_n x = (x_1, \cdots, x_n, 0, \cdots)$, 则 $T_n \in L(l^2)$, 求 $\| T_n \|$.

5. 设 $1 < p = q/(q-1) < \infty, \beta \triangleq \left[\sum_i \left(\sum_j |a_{ij}|^q\right)^{p/q}\right]^{1/p} < \infty, A = [a_{ij}]$，证明 $\|A\|_p \leqslant \beta$.

6. 设 $a \in L^\infty(J), Tu(x) = a(x)u(x)$，则 $T \in L(L^2(J))$，求 $\|T\|$.

7. 设 $Tu(x) = x\int_0^1 u(y)\mathrm{d}y$，则 $T \in L(L^1(J))$ 且 $T \in L(L^\infty(J)), J = [0,1]$，求 $\|T\|_1$ 与 $\|T\|_\infty$.

8. 设 $Tu(x) = \int_0^1 (x-y)u(y)\mathrm{d}y$，求 $\|T\|_1$ 与 $\|T\|_\infty$.

9. 设 $Tu(x) = \int_0^1 u(y)\sin\pi(x-y)\mathrm{d}y$，则 $T \in L(C[0,1])$，求 $\|T\|$.

10. 设 $1 < p = q/(q-1) < \infty, \beta \triangleq \left\{\int_a^b \left[\int_a^b |K(x,y)|^q \mathrm{d}y\right]^{p/q} \mathrm{d}x\right\}^{1/p} < \infty, Tu(x) = \int_a^b K(x,y)u(y)\mathrm{d}y$，证明 $\|T\|_p \leqslant \beta$.

11. 设 $u(x) = \sum x_i/i^\alpha, 0 < \alpha < 1, u \in (l^p)^*$，求 p 的取值范围.

12. 设 $u(x) = \sum x_i/i, u \in (l^2)^*$，求 $\|u\|$.

13. 设 $f(u) = \int_0^\pi u(x)\sin nx\,\mathrm{d}x, f \in L^2[0,\pi]^*$，求 $\|f\|$.

14. 设 $f(u) = \int_0^1 xu(x)\mathrm{d}x + u(1), f \in C[0,1]^*$，求 $\|f\|$.

15. 设 $f(u) = u(1) - u(0), f \in C[0,1]^*$，求 $\|f\|$.

16. 设 $(X, \|\cdot\|)$ 与 $(X, \|\cdot\|')$ 均为 Banach 空间，$\|x\|' \leqslant \mathrm{const}\|x\| (\forall x \in X)$，证明范数 $\|\cdot\|$ 与 $\|\cdot\|'$ 等价.

17. 设 X 是 Banach 空间，$X = A \oplus B$ 是拓扑直和. 证明投影 P_A 与 P_B 均连续.

18. 设 $T \in L(L^2(J)), TC(J) \subset C(J)$，证明 $T \in L(C(J))$.

19. 设 $\{T_n\} \subset L(X,Y), X$ 完备，$\sup_n |v(T_n x)| < \infty (\forall x \in X, \forall v \in Y^*)$，证明 $\sup_n \|T_n\| < \infty$.

20. 设 $T: X \to Y$ 是线性算子，$\forall v \in Y^*$，有 $v \circ T \in X^*$，证明 T 有界.

21. 设 A 是 X 的闭子空间. 证明 $d(x,A) = \sup\{|u(x)|/\|u\| : 0 \neq u \in A^\perp\}$.

22. 设 A 是 Hilbert 空间 H 的子空间，$u \in A^*$，证明 u 在 H 上有惟一保范延拓.

23. 任给 $\varnothing \neq A \subset X$，证明 $A^\perp$ 是 X^* 的闭子空间，$A^\perp = (\overline{\mathrm{span}A})^\perp$.

24. 任给 $\varnothing \neq A \subset X$，证明 $\overline{\mathrm{span}A} = {}^\perp(A^\perp)$.

25. 叙述并证明关于向量值函数的分部积分公式.

26. 给定 $\alpha_i \in \mathbf{K}, x_i \in X (1 \leqslant i \leqslant n)$. 证明：方程组 $\langle u, x_i\rangle = \alpha_i (1 \leqslant i \leqslant n)$ 有解 $u \in X^*$ 的充要条件是

$$\left|\sum \alpha_i\beta_i\right| \leqslant \mathrm{const}\left\|\sum \beta_i x_i\right\| \quad ((\beta_i) \in \mathbf{K}^n).$$

27. 设 A 是实赋范空间 X 中的闭凸集，$A \neq X$. 证明 A 是闭半空间的交.

28. 证明：在 Hilbert 空间中，$x_n \to x \Leftrightarrow x_n \rightharpoonup x$ 且 $\|x_n\| \to \|x\|$.

29. 证明:弱收敛的极限有惟一性.

30. 设在 X 中 $x_n \rightharpoonup x$, 在 X^* 中 $u_n \to u$ (或 $x_n \to x$, $u_n \overset{*}{\rightharpoonup} u$, X 完备), 证明 $u_n(x_n) \to u(x)$.

31. 设 $T: X \to Y$ 是线性算子, $x_n \to x \Rightarrow Tx_n \rightharpoonup Tx$, 证明 T 有界.

32. 证明:若 A 是实赋范空间 X 中的闭凸集, $\{x_n\} \subset A$, $x_n \rightharpoonup x$, 则 $x \in A$.

33. 设 $f_n(u) = \int_a^b u(x)\sin nx\,\mathrm{d}x$, $u \in L^p[a,b]$, $1 \leqslant p < \infty$, 证明 $f_n \overset{*}{\rightharpoonup} 0$.

34. 设 $f_n(u) = \int_a^b u(x)\,|\sin nx|\,\mathrm{d}x$, $f(u) = \dfrac{2}{\pi}\int_a^b u(x)\mathrm{d}x$, $u \in L^p[a,b]$, $1 \leqslant p < \infty$, 证明 $f_n \overset{*}{\rightharpoonup} f$.

35. 设 $T_n u(x) = n\int_x^{x+1/n} u(y)\mathrm{d}y$, $u \in L^1[0,1]$, 在 $[0,1]$ 之外认定 $u = 0$. 证明 $\{T_n\} \subset L(L^1[0,1])$ 强收敛于单位算子.

第三章　谱 论 初 步

在上一章中我们看到，以矩阵比拟有界线性算子，是一个颇富启发性的想法。在本章中，我们要继续运用这种比拟。从线性代数知道，矩阵理论的最精彩且最有价值的部分，莫过于特征值理论。特征值概念源于对线性方程

$$\lambda x - Ax = y$$

的研究，而其结果的应用则远远超出线性方程理论的范围。利用特征值理论所解决的问题可以举出：

(A) 将矩阵化成某种标准形，并将矩阵按其标准形分类；

(B) 求得矩阵的一定分解，并由之得出空间的相应分解；

(C) 求得二次型的标准形，并按标准形进行分类，等。

一个看来过于大胆的设想是：将对矩阵所做到的一切推广到 Banach 空间中去。如果想不受限制地做到这一点，那就期望过高了。值得庆幸的是，在一定限制下，在 Banach 空间中确实能展开一个高度类似的理论，它就是线性算子的谱理论。算子谱论迄今仍然是泛函分析中最优美的部分，它的内容异常丰富，本章的介绍不过是入门而已。

本章中 X 总记一个给定的复 Banach 空间①，且 $X \neq \{0\}$。不过，本章的某些结果亦适用于实 Banach 空间，这种情况可依上下文进行判断。

§3.1　有界线性算子的谱

§3.1.1　算子代数

鉴于我们已提到本章明显的线性代数背景，以关于 $L(X)$ 的一个代数讨论起步是很自然的。我们将处处以矩阵代数作为引导思路的模型。如 §2.1.3 中已指出的(参看命题 2.1.6)，$L(X)$ 是一个复Banach 空间。与一般的有界线性算子空间 $L(X,Y)$ 不同的是，$L(X)$ 对于算子的乘法运算(依式(2.1.3))封闭。这样，$L(X)$ 是一个具有线性运算与乘法运算的代数系统，通常称之为**算子代数**。当 $X = \mathbf{C}^n$ 时，$L(\mathbf{C}^n)$ 可等同于 $\mathbf{C}^{n\times n}$，即 n 阶复矩阵之全体，称之为**矩阵代数**。从纯

---

① 如同在线性代数中所知的，讨论谱或特征值问题不可能避开复数。

代数方面考虑，一般的算子代数 $L(X)$ 具有如同矩阵代数 $\mathbf{C}^{n\times n}$ 一样的性质.这意味着，你几乎不必改动，就可将线性代数中关于矩阵运算的那些规则移用于 $L(X)$ 中的运算.这样，对于任给的 $R,S,T\in L(X)$，$\alpha\in\mathbf{C}$，以下结论成立(参看§2.1.1)：

(i) 结合律：$(RS)T=R(ST)$；因此乘幂 $T^n(n\in\mathbf{N})$ 有确定意义，且

$$T^{m+n}=T^mT^n\qquad(m,n\in\mathbf{N}).$$

(ii) 联系数乘的结合律：$\alpha(ST)=(\alpha S)T=S(\alpha T)$.

(iii) 对加法的分配律：

$$R(S+T)=RS+RT,\quad(R+S)T=RT+ST.$$

(iv) 单位算子 I 是乘法单位元：$IT=TI=T$.

(v) 可逆性：$T:X\to X$ 为同构⇔存在 $A,B\in L(X)$，使得 $AT=I=TB$；必定 $A=B$，称它为 T 的逆，记作 T^{-1}，并称 T 为**可逆算子**①.约定以 $\mathrm{GL}(X)$ 记 $L(X)$ 中的可逆算子之全体.

(vi) 若 $S,T\in\mathrm{GL}(X)$，则 $ST\in\mathrm{GL}(X)$，且

$$(ST)^{-1}=T^{-1}S^{-1},\quad(T^n)^{-1}=(T^{-1})^n.\tag{3.1.1}$$

当 $T\in\mathrm{GL}(X)$ 时约定 $T^{-n}=(T^{-1})^n(n>0)$，$T^0=I$，因而对任何 $k\in\mathbf{Z}$，T^k 有意义.

如同矩阵乘法一样，算子乘法并**不满足**交换律，在作乘法运算时务必小心.若 $T,S\in L(X)$ 满足 $TS=ST$，则说 T 与 S **可换**.两算子可换是很不寻常的事.

至此为止，尚未涉及超出与矩阵代数相当的内容.只有考虑极限运算，才能获得新的知识；而为此必须使用范数.设 $S,T\in L(X)$，$x\in X$，则有(用式(2.1.6))

$$\|STx\|\leqslant\|S\|\|Tx\|\leqslant\|S\|\|T\|\|x\|,$$

由此得出(用式(2.1.9))

$$\|ST\|\leqslant\|S\|\|T\|,\quad\|T^n\|\leqslant\|T\|^n(n\geqslant1).\tag{3.1.2}$$

在一定意义上可以说，前面所述的代数性质(i)～(vi)加上不等式(3.1.2)，就完全描述了算子代数 $L(X)$.因此，不等式(3.1.2)在本章中将起基本作用.利用式(3.1.2)可推出算子乘法的连续性：若在 $L(X)$ 中 $S_n\to S$，$T_n\to T$ ②，则必 $S_nT_n\to ST$，这由如下简单推理得出：

---

① 若 $T\in L(X)$ 是单射，则存在逆算子 $T^{-1}:R(T)\to X$.不过，本章中说到可逆算子 T 时，严格限定指以下情况：T^{-1} 存在且定义于整个空间 X 上.

② 在未加说明时，本章用到的算子序列收敛都指一致收敛.

$$\begin{aligned}\|S_nT_n - ST\| &\leqslant \|S_nT_n - S_nT\| + \|S_nT - ST\| \\ &\leqslant \|S_n\|\|T_n - T\| + \|S_n - S\|\|T\| \quad (\text{用式}(3.1.2)) \\ &\leqslant \text{const}(\|T_n - T\| + \|S_n - S\|) \to 0 \quad (n \to \infty).\end{aligned}$$

在§1.1.2中我们看到,在任何 Banach 空间中可以运用无穷级数.而在具有乘法运算的算子代数中,则可进而考虑**算子幂级数**:

$$f(T) = \sum_{n=0}^{\infty} \alpha_n T^n = \alpha_0 I + \alpha_1 T + \cdots + \alpha_n T^n + \cdots, \tag{3.1.3}$$

其中系数 $\alpha_n \in \mathbf{C}(n \geqslant 0)$.给出级数(3.1.3)的收敛条件,将是本节的目标之一.一个不那么精细但很简单的收敛判别法是:若通常幂级数

$$f(\lambda) = \sum_{n=0}^{\infty} \alpha_n \lambda^n \tag{3.1.4}$$

有收敛半径 R,则当 $T \in L(X)$,$\|T\| < R$ 时级数(3.1.3)绝对收敛:

$$\sum \|\alpha_n T^n\| \leqslant \sum |\alpha_n| \|T\|^n < \infty.$$

本章将多次用到如下基本引理.

引理 3.1.1 设 $T \in L(X)$,则

$$(I - T)^{-1} = \sum_{n=0}^{\infty} T^n, \tag{3.1.5}$$

只要其右端级数收敛.特别,当 $\|T\| < 1$ 时(3.1.5)必成立.

证 设式(3.1.5)右端级数收敛于 $S \in L(X)$,则

$$\begin{aligned}(I - T)S &= S - TS \\ &= \sum_{n=0}^{\infty} T^n - \sum_{n=1}^{\infty} T^n = I.\end{aligned}$$

同理 $S(I - T) = I$,因此 $S = (I - T)^{-1}$. □

读者想必已看出,式(3.1.5)正是展开式

$$\frac{1}{1-\lambda} = \sum_{n=0}^{\infty} \lambda^n \qquad (|\lambda| < 1),$$

的推广.不妨就称式(3.1.5)为算子 $(I - T)^{-1}$ 的**二项式级数**.

通常依如下形式应用引理3.1.1:若 $T, S \in L(X)$,T 可逆,则

$$(T + S)^{-1} = T^{-1}(I + ST^{-1})^{-1} = \sum_{n=0}^{\infty} T^{-1}(-ST^{-1})^n, \tag{3.1.6}$$

只要其右端之级数收敛;特别,当 $\|S\|$ 适当小时式(3.1.6)必成立.

§3.1.2　谱与谱半径

对于 $A \in \mathbf{C}^{n\times n}, \lambda \in \mathbf{C}$, 熟知以下关系成立:

$$\lambda \text{ 是} A \text{ 的特征值} \Leftrightarrow \det(\lambda I - A) \neq 0$$

$$\Leftrightarrow \lambda I - A \text{ 不可逆}.$$

这提示出以下定义.

定义 3.1.2　设 $T \in L(X)$,

(i) 若 $\lambda \in \mathbf{C}, \lambda I - T$ 不可逆,即 $\lambda I - T \in \mathrm{GL}(X)$, 则称 λ 为 T 的**谱值**. 以 $\sigma(T)$ 记 T 的谱值之全体,称它为 T 的**谱**;称

$$r_\sigma(T) = \sup_{\lambda \in \sigma(T)} |\lambda| \tag{3.1.7}$$

为 T 的**谱半径**,它就是以原点为中心且包含 $\sigma(T)$ 的最小的圆的半径.

(ii) 令 $\rho(T) = \mathbf{C} \setminus \sigma(T)$, 称任何 $\lambda \in \rho(T)$ 为 T 的**正则值**;称

$$R(\lambda, T) = (\lambda I - T)^{-1} \quad (\lambda \in \rho(T)) \tag{3.1.8}$$

为**预解式**. 当不致误解时,也将 $R(\lambda, T)$ 写作 $R(\lambda)$ 或 R_λ.

(iii) 若 $\lambda \in \mathbf{C}$, 存在 $x \neq 0$, 使得 $Tx = \lambda x$ (这相当于 $x \in N(\lambda I - T)$),则称 λ 为 T 的**特征值**,并称 x 为 T 关于 λ 的**特征向量**,称 $N(\lambda I - T)$ 为 T 关于特征值 λ 的**特征子空间**. 以 $\sigma_p(T)$ 记 T 的特征值之全体,称它为 T 的**点谱**.

因 $N(\lambda I - T) \neq \{0\} \Rightarrow \lambda I - T$ 不可逆,故 $\sigma_p(T) \subset \sigma(T)$. 若 $T \in L(X)$, $\dim X < \infty$, 则 $N(\lambda I - T) - \{0\} \Leftrightarrow \lambda I - T$ 可逆,因而 $\sigma_p(T) - \sigma(T)$. 若 $\dim X = \infty$, 则可能 $\sigma_p(T) \neq \sigma(T)$, 即谱值未必是特征值. 这一事实提醒我们:不可不加思索地将矩阵特征值理论中的任何结论用于算子谱论!

对于 $A \in \mathbf{C}^{n\times n}$, 原则上我们能计算 A 的全部特征值(它至多 n 个且由一 n 次代数方程解出). 对于一般的 $T \in L(X)$, 并无计算其谱值的一般方法. 我们甚至不知道, $\sigma(T)$ 是有限集还是无限集. 因此,关于 $\sigma(T)$ 性质的任何一般性结论均属可贵. 以下定理则是一个特别令人满意的结果.

定理 3.1.3(Gelfand 定理)　设 $T \in L(X)$, 则 $\sigma(T)$ 是非空紧集,且成立**谱半径公式**:

$$r_\sigma(T) = \lim_n \|T^n\|^{1/n}. \tag{3.1.9}$$

证明有点长,在初次阅读时可以跳过它.

证　当 $T = 0$ 时显然有 $\sigma(T) = \{0\}$, 式(3.1.9)平凡地成立. 下面设 $T \neq 0$,

因而 $\|T\| > 0$.

(i) 证 $\lim_n \|T^n\|^{1/n} = \inf_n \|T^n\|^{1/n} \triangleq \beta$. 显然

$$\underline{\lim_n} \|T^n\|^{1/n} \geqslant \beta,$$

故只要证

$$\overline{\lim_n} \|T^n\|^{1/n} \leqslant \beta.$$

为此,又只要证

$$\overline{\lim_n} \|T^n\|^{1/n} \leqslant \|T^m\|^{1/m}, \quad m = 1,2,\cdots$$

取定 $m \in \mathbf{N}$. $\forall n \in \mathbf{N}$, 有分解 $n = pm + q$, 其中 $p,q \in \mathbf{Z}_+$, $0 \leqslant q < m$. 重复应用不等式(3.1.2)得

$$\|T^n\| \leqslant \|T^{pm}\| \|T^q\| \leqslant \|T^m\|^p \|T\|^q,$$

于是

$$\begin{aligned}\overline{\lim_n} \|T^n\|^{1/n} &\leqslant \overline{\lim_n} \|T^m\|^{p/n} \|T\|^{q/n} \\ &= \overline{\lim_n} \|T^m\|^{p/(pm+q)} = \|T^m\|^{1/m},\end{aligned}$$

如所要证.

(ii) 证 $r_\sigma(T) \leqslant \beta$. 为此只要证:若 $\lambda \in \mathbf{C}$, $|\lambda| > \beta$, 则 $\lambda \in \rho(T)$, 即算子 $\lambda I - T$ 可逆,这相当于 $I - \lambda^{-1}T$ 可逆.形式地展开 $(I - \lambda^{-1}T)^{-1}$:

$$(I - \lambda^{-1}T)^{-1} = \sum_{n=0}^{\infty} \lambda^{-n}T^n.$$

由引理 3.1.1,只要证明上式右端级数收敛,而这可用熟知的 Cauchy 判别法:由

$$\lim_n \|\lambda^{-n}T^n\|^{1/n} = \beta/|\lambda| < 1$$

推出级数 $\sum \lambda^{-n}T^n$ 绝对收敛. 于是当 $|\lambda| > \beta$ 时有

$$R(\lambda, T) = (\lambda I - T)^{-1} = \sum_{n=0}^{\infty} \lambda^{-n-1}T^n. \tag{3.1.10}$$

(iii) 证 $\sigma(T)$ 为闭集,为此只要证 $\rho(T)$ 为开集.取定 $\lambda_0 \in \rho(T)$. 若 $\lambda \in \mathbf{C}$, $|\lambda - \lambda_0|$ 充分小,则

$$\lambda I - T = (\lambda_0 I - T) + (\lambda - \lambda_0)I$$

必可逆,因而 $\lambda \in \rho(T)$, 这得出 $\rho(T)$ 为开集.令 $R_0 = R(\lambda_0, T)$, 分别以 R_0^{-1} 与

$(\lambda-\lambda_0)I$ 代 T 与 S 应用式(3.1.6),得

$$R(\lambda,T)=\sum_{n=0}^{\infty}(\lambda_0-\lambda)^n R_0^{n+1}. \tag{3.1.11}$$

(iv) 证 $\sigma(T)\neq\varnothing$. 用反证法:设 $\sigma(T)=\varnothing$, 则 $R_\lambda=R(\lambda,T)$ 对任何 $\lambda\in\mathbf{C}$ 有定义. 任给 $x\in X, u\in X^*$, 定义复函数

$$f(\lambda)=u(R_\lambda x)\quad(\lambda\in\mathbf{C}).$$

局部地,可将 R_λ 表为式(3.1.11),因而

$$f(\lambda)=\sum_{n=0}^{\infty}u(R_0^{n+1}x)(\lambda_0-\lambda)^n,$$

这表明 $f(\lambda)$ 处处解析,因而是一个整函数. 当 $|\lambda|>\beta$ 时,利用(10)可得

$$f(\lambda)=\sum_{n=0}^{\infty}\lambda^{-n-1}u(T^n x), \tag{3.1.12}$$

这表明 $f(\infty)=0$. 于是由复变函数论中的 Liouville 定理有 $f(\lambda)\equiv 0$. 由 x,u 的任意性得 $R_\lambda\equiv 0$ (用推论 2.4.7),但这与 R_λ 可逆矛盾.

由已证的(ii)~(iv),知 $\sigma(T)$ 为非空紧集.

(v) 证 $r_\sigma(T)\geqslant\beta$ (从而 $r_\sigma(T)=\beta$, 即式(3.1.9)成立). 为此只需证,对任给 $\alpha>r_\sigma(T)$, 有 $\beta\leqslant\alpha$. 任给 $x\in X, u\in X^*$, $f(\lambda)=u(R_\lambda x)$ 在 $|\lambda|>r_\sigma(T)$ 内解析,因而其 Laurent 级数式(3.1.12)在 $\lambda=\alpha$ 时绝对收敛,于是

$$\sup_n|u(\alpha^{-n}T^n x)|<\infty.$$

因 x,u 是任取的,两次应用一致有界原理得出

$$\sup_n\|\alpha^{-n}T^n\|<\infty.$$

于是

$$1\geqslant\lim_n\|\alpha^{-n}T^n\|^{1/n}=\beta/\alpha,$$

即 $\beta\leqslant\alpha$, 如所要证. □

定理 3.1.3 断定 $\sigma(T)$ 是非空紧集,固然是重要的,但并不算是很精细的结论. 实际上,容易举例说明,任何非空紧集 $\sigma\subset\mathbf{C}$ 都可能成为某个有界线性算子的谱(见习题 7),定理 3.1.3 中真正惊人的结论是谱半径公式(3.1.9),如此准确的定量关系在泛函分析中殊不多见. 如果考虑到, $\sigma(T)$ 与 X 中的等价范数的选择无关,而式(3.1.9)右端似乎依赖于范数的选择,就更显示出公式(3.1.9)的深刻性.

由公式(3.1.9)直接推出一个很简单的估计:

$$r_\sigma(T) \leqslant \| T \| . \tag{3.1.13}$$

不等式(3.1.13)用起来很简便，只是过于粗略；有例子表明，即使 $\| T \|$ 很大，亦有可能 $r_\sigma(T) = 0$[①].

§3.1.3 某些应用

Gelfand 定理应用极广，此处仅考虑一些初步的例子. 首先，谱半径公式为算子幂级数提供了方便的收敛判别法.

定理 3.1.4 设幂级数 $\sum \alpha_n \lambda^n$ 的收敛半径为 R，$T \in L(X)$.

(i) 若 $r_\sigma(T) < R$，则级数 $\sum \alpha_n T^n$ 绝对收敛.

(ii) 若 $r_\sigma(T) > R$，则级数 $\sum \alpha_n T^n$ 发散.

证 从微积分学知，收敛半径 R 有计算公式

$$1/R = \overline{\lim_n} \mid \alpha_n \mid^{1/n}.$$

这与谱半径公式相结合得出：

$$\overline{\lim_n} \| \alpha_n T^n \|^{1/n} = \overline{\lim_n} \mid \alpha_n \mid^{1/n} \lim_n \| T^n \|^{1/n} = r_\sigma(T)/R. \tag{3.1.14}$$

(i) 若 $r_\sigma(T) < R$，则由式(3.1.14)有 $\overline{\lim_n} \| \alpha_n T^n \|^{1/n} < 1$，于是由 Cauchy 判别法知级数 $\sum \alpha_n T^n$ 绝对收敛.

(ii) 若级数 $\sum \alpha_n T^n$ 收敛，则 $\alpha_n T^n \to 0$，于是由式(3.1.14)有

$$r_\sigma(T)/R = \overline{\lim_n} \| \alpha_n T^n \|^{1/n} \leqslant 1,$$

即 $r_\sigma(T) \leqslant R$. 可见当 $r_\sigma(T) > R$ 时级数 $\sum \alpha_n T^n$ 必发散. □

这样，除了 $r_\sigma(T) = R$ 的情况(这意味着在圆周 $\mid \lambda \mid = R$ 上存在 T 的谱值)以外，定理 3.1.4 完全解决了算子幂级数的收敛判别问题. 若 $r_\sigma(T) = R$，则由通常幂级数的情况就知道，级数 $\sum \alpha_n T^n$ 既可能收敛，也可以发散.

特别将定理 3.1.4 用于级数 $\sum T^n$ 得出：只要 $r_\sigma(T) < 1$，就有展开式(3.1.5)成立，且(3.1.5)右端之级数绝对收敛. 这是一个远比引理 3.1.1 更精细的结论. 将这一结论用到线性方程

$$x = Tx + y \tag{3.1.15}$$

---

① 例如取 $A \in \mathbf{C}^{n \times n}$，使 $A \neq 0, A^k = 0$，令 $T = \alpha A$，则 $\| T \|$ 可以随 $\mid \alpha \mid$ 任意变大，但保持 $r_\sigma(T) = 0$.

($y\in X$ 是已知的)得出;若 $T\in L(X)$, $r_\sigma(T)<1$, 则方程(3.1.15)有惟一解 x, 且

$$x=(I-T)^{-1}y=\sum_{n=0}^{\infty}T^n y,$$

上式右端之级数绝对收敛,通常称它为 **Neumann 级数**.

顺便指出,若令 $Fx=Tx+y$, 取 $x_0=y$, 归纳出定义

$$x_n=Fx_{n-1}=F^n x_0\quad(n\geqslant 1),\tag{3.1.16}$$

则 $x_n=\sum_0^n T^k y$. 因此由式(3.1.16)表出的 x 正好是迭代序列 $\{x_n\}$ 的极限. 这就表明,只要 $r_\sigma(T)<1$, 就可用迭代法求方程(3.1.15)的解. 在§4.2 中,我们将在更一般的条件下讨论迭代法.

下面是一个应用以上结论的具体例子.

例 3.1.5 设 $J=[0,1]$, 给定 $b(\cdot)\in C(J)$, 考虑积分方程

$$u(x)=\int_0^x u(y)\mathrm{d}y+b(x).\tag{3.1.17}$$

若令

$$Tu(x)=\int_0^x u(y)\mathrm{d}y\quad(x\in J),$$

则可将方程(3.1.17)改写成

$$u=Tu+b,\tag*{(3.1.17)$'$}$$

这正是一个形如式(3.1.15)的线性方程. 显然 $T\in L(C(J))$. 用归纳法可验证

$$\begin{aligned}T^n u(x)&=\int_0^x\mathrm{d}x_1\int_0^{x_1}\mathrm{d}x_2\cdots\int_0^{x_{n-1}}u(x_n)\mathrm{d}x_n\\&=\frac{1}{(n-1)!}\int_0^x(x-y)^{n-1}u(y)\mathrm{d}y\quad(x\in J,n\geqslant 1),\end{aligned}\tag{3.1.18}$$

因而

$$|T^n u(x)|\leqslant\frac{\|u\|_0}{(n-1)!}\int_0^1(1-y)^{n-1}\mathrm{d}y=\frac{\|u\|_0}{n!},$$

这得出 $\|T^n\|\leqslant 1/n!$, 于是

$$r_\sigma(T)\leqslant\lim_n(n!)^{-1/n}=0.$$

因此,方程(3.1.17)$'$(或方程(3.1.17))有惟一解 $u\in C(J)$, 且

$$u(x) = \sum_{n=0}^{\infty} T^n b(x)$$

$$= b(x) + \sum_{n=1}^{\infty} \frac{1}{(n-1)!}\int_0^x (x-y)^{n-1} b(y)\mathrm{d}y \quad (\text{用}(18))$$

$$= b(x) + \int_0^x \sum_{n=0}^{\infty} \frac{(x-y)^n}{n!} b(y)\mathrm{d}y$$

$$= b(x) + \int_0^x \mathrm{e}^{x-y} b(y)\mathrm{d}y.$$

§3.2 算子函数

§3.2.1 解析扩张

定理 3.1.4 的一个最有意义的推论是:若

$$f(\lambda) = \sum_{n=0}^{\infty} a_n (\lambda - \lambda_0)^n \tag{3.2.1}$$

是圆

$$D_r(\lambda_0) \triangleq \{\lambda \in \mathbf{C} : |\lambda - \lambda_0| < r\} \tag{3.2.2}$$

内的复解析函数,则当 $T \in L(X), r_\sigma(T - \lambda_0 I) < r$ 时,

$$f(T) = \sum_{n=0}^{\infty} a_n (T - \lambda_0 I)^n \tag{3.2.3}$$

有意义,且式(3.2.3)右端之级数绝对收敛. 依定义 3.1.2(i)易直接验证 $\sigma(T - \lambda_0 I) = \sigma(T) - \lambda_0 = \{\lambda - \lambda_0 : \lambda \in \sigma(T)\}$, 于是

$$r_\sigma(T - \lambda_0 I) < r \Leftrightarrow \sigma(T - \lambda_0 I) \subset D_r(0)$$

$$\Leftrightarrow \sigma(T) \subset \lambda_0 + D_r(0) = D_r(\lambda_0).$$

因此,式(3.2.3)表示一个定义于集 $\{T \in L(X) : \sigma(T) \subset D_r(\lambda_0)\}$ 上的算子函数 $f(T)$. 自然,可将 $f(T)$ 看作复解析函数 $f(\lambda)$ 的某种扩张. 特别,熟知的初等函数都可适当地扩张为算子函数. 例如,对数函数

$$\ln \lambda = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}(\lambda - 1)^n}{n} \qquad (\lambda \in D_1(1))$$

可扩张为集 $\{T \in L(X) : \sigma(T) \subset D_1(1)\}$ 上的**算子对数函数**

$$\ln T = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}(T-I)^n}{n}.$$

类似地,还可定义算子的指数函数 e^T、正弦函数 $\sin T$,等等.这样,我们似乎已经轻而易举地从复解析函数过渡到了**算子解析函数**.

这实在是很引人的.不过,我们立即会发现以下问题:

(A) 幂级数仅能表达圆域内的解析函数.对任意开集 $\Omega(\subset \mathbf{C})$ 内的解析函数 $f(\lambda)$ ①及满足 $\sigma(T) \subset \Omega$ 的 $T \in L(X)$,应如何定义 $f(T)$?

(B) 算子函数 $f(T)$ 真正类似于其原型 $f(\lambda)$ 吗?例如,如果不能使用恒等式

$$e^T e^{-T} = I, \quad e^{\ln T} = T,$$

我们能将 e^T 与 $\ln T$ 认作指数函数与对数函数吗?根本的问题是,$f(T)$ 能继承函数 $f(\lambda)$ 的哪些性质?

(C) 函数 $f(T)$ 仅只是 $f(\lambda)$ 的形式扩张,还是有某些不可缺少的实质性应用?

幸而,对这些问题已有令人充分满意的解答.不过,要在短短的一节中完全地陈述这些解答并非易事.我们将主要强调展开理论所循的思路,而不拘泥于某些推导细节.

首先,考虑任意复解析函数的扩张问题.取定非空开集 $\Omega \subset \mathbf{C}$.为行文简便起见,以 $H(\Omega)$ 记 Ω 内的复解析函数之全体,令

$$D_\Omega = \{T \in L(X) : \sigma(T) \subset \Omega\}. \tag{3.2.4}$$

设 $f(\lambda) \in H(\Omega)$, $T \in D_\Omega$,今探求 $f(T)$ 的一个合理定义.因未必有某个圆 $D_r(\lambda_0)$,使得

$$\sigma(T) \subset D_r(\lambda_0) \subset \Omega,$$

形如式(3.2.3)的定义式一般不再有效.现在你该回想到,复解析函数不仅可表为幂级数,而且可表为积分,即有如下的 **Cauchy 公式**表示:

$$f(\lambda) = \frac{1}{2\pi i}\int_L f(\tau)(\tau-\lambda)^{-1}d\tau, \tag{3.2.5}$$

其中 L 是 Ω 内任一围绕 λ 的简单闭曲线(为行文简便,下面称之为**围道**,且假定沿其正向行进时,保持 λ 所在区域在左边).观察了式(3.2.1)与式(3.2.3)的关系之后,你大概会猜想,式(3.2.5)正好推广为一个算子函数的表达式:

① 因开集不必是区域,准确地说,这样的 $f(\lambda)$ 应称为**局部解析函数**.

$$f(T) = \frac{1}{2\pi i}\int_L f(\tau)(\tau I - T)^{-1}d\tau. \tag{3.2.6}$$

可惜,我们还没有关于算子值函数的积分概念.下面就来补叙有关定义.

设 L 是复平面上任一可求长曲线, $T(\tau)$ 是定义于 L 上而取值于 $L(X)$ 中的函数(称为**算子值函数**),则可用通常的"分割、求和、取极限"的方式定义 $T(\tau)$ 沿 L 的积分:

$$\int_L T(\tau)d\tau = \lim_{\max|\Delta\tau_i|\to 0}\sum_{i=1}^{n} T(\xi_i)\Delta\tau_i, \tag{3.2.7}$$

其中 $\tau_0,\tau_1,\cdots,\tau_n$ 为 L 上顺次排列的分点, τ_0 与 τ_n 分别为 L 的起点与终点, ξ_i 是 L 上介于 τ_{i-1} 与 τ_i 之间的任一点, $\Delta\tau_i = \tau_i - \tau_{i-1}(1\leqslant i\leqslant n)$. 用类似于复变函数中的方法不难证明,当 $T(\tau)$ 对 τ 连续时,积分(3.2.7)必存在.利用定义式(3.2.7)直接推出,对任给 $u\in X^*$ 与 $x\in X$ 有(参照式(2.4.9)、(2.4.10))

$$\langle u,\int_L T(\tau)d\tau\cdot x\rangle = \int_L\langle u,T(\tau)x\rangle d\tau. \tag{3.2.8}$$

作了上述准备之后,现在已可陈述本节的主要定义.

定义 3.2.1 任给 $f(\lambda)\in H(\Omega)$ 与 $T\in D_\Omega$,取 Ω 内任一围绕 $\sigma(T)$ 的围道 L,依式(3.2.6)定义 $f(T)$,则得到一个从 D_Ω 到 $L(X)$ 的函数 $f(T)$,称它为 $f(\lambda)$ 的**解析扩张**,或简称为扩张.

对于以上定义有几个明显的疑问需要澄清.

(i) 要说明式(3.2.6)右端的积分必定存在.首先,因 $\sigma(T)$ 是开集 Ω 的紧子集, Ω 内必存在围绕 $\sigma(T)$ 的围道[①].其次,式(3.1.11)显然蕴涵了 $R(\tau,T)$ 对 τ 的连续性.因此,(3.2.6)式中的被积函数 $f(\tau)R(\tau,T)$ 对 τ 在 L 上连续,这就保证了积分存在.

(ii) 要说明(3.2.6)右端积分不依赖于 L 的选择.为此,又只要说明,对任给 $u\in X^*$, $x\in X$,通常的复积分

$$\langle u,f(T)x\rangle = \frac{1}{2\pi i}\int_L f(\tau)\langle u,R(\tau,T)x\rangle d\tau \tag{3.2.9}$$

不依赖于 L 的选择(你看出这里用了(3.2.8)与推论 2.4.7).而由复变函数知识,只要说明式(3.2.9)中的被积函数解析就行了.回顾一下定理 3.1.3 的证明你会看出,我们已经指出了 $\langle u,R(\tau,T)x\rangle$ 关于 τ 的解析性.

(iii) 定义式(3.2.3)与(3.2.6)(当它们都可使用时)是一致的.事实上,若

---

① 这一事实的严格证明并不像读者初看起来那样简单.

$\sigma(T) \subset D_r(\lambda_0) \subset \Omega$，取 $0 < \rho < r$，使 $\sigma(T) \subset D_\rho(\lambda_0)$；令 $L = \{\lambda \in \mathbf{C} : |\lambda - \lambda_0| = \rho\}$，则

$$\begin{aligned}
&\frac{1}{2\pi i}\int_L f(\lambda)(\lambda I - T)^{-1}d\lambda \\
&= \frac{1}{2\pi i}\int_L f(\lambda)[(\lambda - \lambda_0)I + \lambda_0 I - T]^{-1}d\lambda \\
&= \frac{1}{2\pi i}\int_L f(\lambda)\sum_{n=0}^{\infty}\frac{(T - \lambda_0 I)^n}{(\lambda - \lambda_0)^{n+1}}d\lambda \qquad (\text{用式}(3.1.6)) \\
&= \sum_{n=0}^{\infty}\left[\frac{1}{2\pi i}\int_L \frac{f(\lambda)}{(\lambda - \lambda_0)^{n+1}}d\lambda\right](T - \lambda_0 I)^n \\
&= \sum_{n=0}^{\infty}\frac{f^{(n)}(\lambda_0)}{n!}(T - \lambda_0 I)^n, \qquad (\text{用 Cauchy 公式})
\end{aligned}$$

其中的 Taylor 系数 $f^{(n)}(\lambda_0)/n!$ 正好是式(3)中的系数 α_n.

(iv) 现在说明：$f(T)$ 确是 $f(\lambda)$ 的扩张. 首先，

$$\mathbf{C} \to L(X), \quad \lambda \to \lambda I$$

显然是一等距嵌入，且此嵌入保持乘积运算. 因此，不妨认为 $\mathbf{C} \subset L(X)$，即等同 λ 与 λI. 显然 $\sigma(\lambda I) = \{\lambda\}$，因此可以认为 $\Omega \subset D_\Omega$. $\forall \lambda \in \Omega$，在 Ω 内取一围绕 λ 的围道 L，则

$$\begin{aligned}
f(\lambda I) &= \frac{1}{2\pi i}\int_L f(\tau)(\tau I - \lambda I)^{-1}d\tau \qquad (\text{用}(3.2.6)) \\
&= \left[\frac{1}{2\pi i}\int_L \frac{f(\tau)}{\tau - \lambda}d\tau\right]I = f(\lambda)I, \quad (\text{用 Cauchy 公式})
\end{aligned}$$

可见 $f(\lambda I)$ 与 $f(\lambda)$ 一致.

这样，本节开头提出的问题(A)已告解决.

§3.2.2 解析扩张的性质

现在考虑关于解析扩张的最重要的问题：$f(T)$ 继承了 $f(\lambda)$ 的哪些性质？关注的重点是运算性质：复解析函数之间的一定运算是否恰好对应于算子解析函数之间的相应运算？形式地说，如果有一个复解析函数的恒等式

$$\varphi(\lambda) = F(f(\lambda), g(\lambda), \cdots, h(\lambda)) \quad (\lambda \in \Omega),$$

其中右端由 $f(\lambda)$，$g(\lambda)$ 等经加法、乘法及复合构成，那么问题是：相应的算子函数

恒等式

$$\varphi(T) = F(f(T), g(T), \cdots, h(T)) \quad (T \in D_\Omega)$$

是否成立？如果回答是肯定的，我们就可以畅通无阻地使用如下恒等式：

$$\sin^2 T + \cos^2 T = I;$$

$$\cos T = \sin(2^{-1}\pi I - T);$$

$$e^{\ln T} = T,$$

其中 $T \in L(X)$，最后一式要求 $\sigma(T) \subset D_1(1)$. 这样一来，对于由扩张而来的算子函数在运用中的巨大方便，你就有一定印象了. 下面的定理正是对此提供了保证，它是本节的中心结果.

定理 3.2.2 设 $f(\lambda), g(\lambda) \in H(\Omega), h(\lambda) \in H(\Omega'), f(\Omega) \subset \Omega' \subset \mathbf{C}, T \in D_\Omega$，则

$$(f+g)(T) = f(T) + g(T); \tag{3.2.10}$$

$$(fg)(T) = f(T)g(T); \tag{3.2.11}$$

$$(h \circ f)(T) = h(f(T)). \tag{3.2.12}$$

证 式(3.2.10)是明显的，只需证式(3.2.11)与(3.2.12). 记 $R(\lambda) = R(\lambda, T)$.

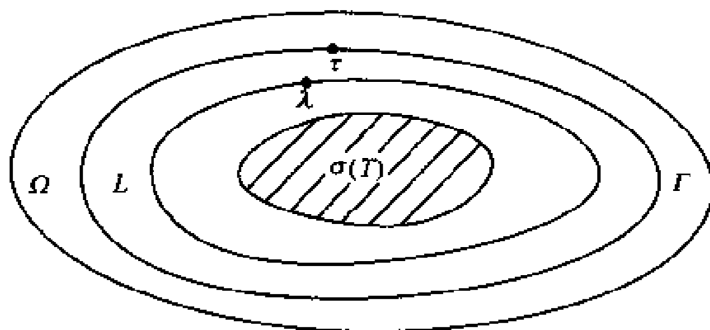


图 3-1

证(3.2.11)式. 选取围道 L 与 Γ 如图 3-1，则

$$\begin{aligned} f(T)g(T) &= \frac{1}{(2\pi i)^2}\int_L f(\lambda)R(\lambda)\mathrm{d}\lambda \int_\Gamma g(\tau)R(\tau)\mathrm{d}\tau \\ &= \frac{1}{(2\pi i)^2}\int_L f(\lambda)\mathrm{d}\lambda \int_\Gamma g(\tau)R(\lambda)R(\tau)\mathrm{d}\tau \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{(2\pi i)^2}\int_L f(\lambda)d\lambda\int_\Gamma g(\tau)\,\frac{R(\lambda)-R(\tau)}{\tau-\lambda}d\tau \\
&= \frac{1}{(2\pi i)^2}\int_L f(\lambda)R(\lambda)d\lambda\int_\Gamma \frac{g(\tau)}{\tau-\lambda}d\tau \\
&\quad + \frac{1}{(2\pi i)^2}\int_\Gamma g(\tau)R(\tau)d\tau\int_L \frac{f(\lambda)}{\lambda-\tau}d\lambda \quad (\text{积分互换}) \\
&= \frac{1}{(2\pi i)^2}\int_L f(\lambda)g(\lambda)R(\lambda)d\lambda \qquad (\text{用 Cauchy 公式}) \\
&= (fg)(T),
\end{aligned}
$$

这表明式(3.2.11)成立.以上演算中用到一个关键恒等式:

$$
R(\lambda)R(\tau) = \frac{R(\lambda)-R(\tau)}{\tau-\lambda}, \tag{3.2.13}
$$

其验证是直接的(请验证).式(3.2.13)可与以下初等恒等式相对照:

$$
\frac{1}{(\lambda-t)(\tau-t)} = \frac{1}{\tau-\lambda}\left(\frac{1}{\lambda-t}-\frac{1}{\tau-t}\right).
$$

为证式(3.2.12),首先对已证的式(3.2.11)作些解释.式(3.2.11)的最重要的推论是:$f(T)$ 与 $g(T)$ 总是可换的:

$$
f(T)g(T) = (fg)(T) = g(T)f(T);
$$

若 $f(\lambda)\neq 0(\forall\lambda\in\Omega)$,则从 $f\cdot(1/f)=1$ 得出

$$
(1/f)(T) = [f(T)]^{-1}. \tag{3.2.14}
$$

以 $\tau - f(\lambda)$ 代 $f(\lambda)$ 用式(3.2.14)得

$$
[\tau I - f(T)]^{-1} = \frac{1}{2\pi i}\int_L \frac{R(\lambda,T)}{\tau-f(\lambda)}d\lambda. \tag{3.2.15}
$$

证(3.2.12)式.因 $\sigma(f(T)) = f(\sigma(T))$(此处提前应用定理3.2.3,注意定理3.2.3之证用不到(3.2.12)式),故 $\sigma(f(T))$ 是 Ω' 的紧子集,因此可在 Ω' 内取一围绕 $\sigma(f(T))$ 的围道 C,不妨设 C 亦围绕 $f(L)$,L 是 Ω 内围绕 $\sigma(T)$ 的围道.于是

$$
\begin{aligned}
h(f(T)) &= \frac{1}{2\pi i}\int_C h(\tau)[\tau I - f(T)]^{-1}d\tau \\
&= \frac{1}{(2\pi i)^2}\int_C h(\tau)d\tau\int_L \frac{R(\lambda,T)}{\tau-f(\lambda)}d\lambda \qquad (\text{用式(3.2.15)}) \\
&= \frac{1}{(2\pi i)^2}\int_L R(\lambda,T)d\lambda\int_C \frac{h(\tau)}{\tau-f(\lambda)}d\tau \quad (\text{积分互换})
\end{aligned}
$$

$$= \frac{1}{2\pi i}\int_l h(f(\lambda))R(\lambda, T)d\lambda \qquad \text{(用 Cauchy 公式)}$$

$$= (h \circ f)(T),$$

如所要证. □

定理 3.2.3(谱映射定理) 设 $f(\lambda) \in H(\Omega)$, $T \in D_\Omega$, 则

$$\sigma(f(T)) = f(\sigma(T)). \tag{3.2.16}$$

证 等式(3.2.16)相当于 $\forall \alpha \in \mathbf{C}$, 有

$$f(\alpha) \overline{\in} \sigma(f(T)) \to \alpha \overline{\in} \sigma(T);$$

$$\alpha \overline{\in} f(\sigma(T)) \Rightarrow \alpha \overline{\in} \sigma(f(T)).$$

今分别证之. 首先设 $f(\alpha) \overline{\in} \sigma(f(T))$, 则 $f(\alpha)I - f(T)$ 可逆, 今要证 $\alpha I - T$ 可逆. 为此, 作 Ω 内的解析函数

$$\varphi(\lambda) = \frac{f(\alpha) - f(\lambda)}{\alpha - \lambda},$$

注意 $\varphi(\lambda)$ 在 $\lambda = \alpha$ 亦解析. 由 $f(\alpha) - f(\lambda) = \varphi(\lambda)(\alpha - \lambda)$ 得

$$f(\alpha)I - f(T) = \varphi(T)(\alpha I - T),$$

故

$$I = [f(\alpha)I - f(T)]^{-1}\varphi(T)(\alpha I - T)$$

$$= (\alpha I - T)\varphi(T)[f(\alpha)I - T]^{-1},$$

这表明 $\alpha I - T$ 可逆, 如所要证.

其次设 $\alpha \overline{\in} f(\sigma(T))$, 则 $g(\lambda) \triangleq [\alpha - f(\lambda)]^{-1}$ 在紧集 $\sigma(T)$ 上有限, 因而必在包含 $\sigma(T)$ 的某个开集 U 内解析. 于是

$$g(T) = [\alpha I - f(T)]^{-1},$$

可见 $\alpha \overline{\in} \sigma(f(T))$, 如所要证. □

由定理 3.2.3 直接推出, 若 $\lambda \in \sigma(T)$, 则 $\lambda^n \in \sigma(T^n)$; 当 T 可逆时 $\lambda^{-1} \in \sigma(T^{-1})$. 在 $\dim X < \infty$ 的情况下, 以上结果在线性代数中是熟知的. 定理 3.2.3 则将这类结果推广到了很一般的情况.

§3.2.3 谱分解

本节开头对解析扩张所提出的问题(A)~(C), 前两个已获得满意解答, 这是你在前两段已看到的. 至于问题(C), 就不可能详尽讨论了. 算子函数的应用及于

各个领域,此处至多以最初步的例子加以说明.从逻辑上说,算子函数的应用可分为两类.其一涉及特定的算子函数,例如指数函数,这当然是人们便于类比,因而容易接受的,但实际上其应用很受局限;其二是一旦问题需要,就依具体条件临时构成所需的算子函数.后一种应用并不涉及特定的算子函数,因而具有更大的灵活性,它正是应用算子函数的主要形式.你可能已注意到,在证明定理 3.2.3 时,实际上已在这种形式上应用了算子函数.下面考虑的谱分解问题,则是成功地应用算子函数的更具说服力的例证.

定理 3.2.4(谱分解定理) 设 $T \in L(X)$,$\sigma(T) = \bigcup_1^n \sigma_i$,$n \geqslant 2$,$\sigma_i$ 是互不相交的非空闭集,则存在 X 的拓扑直和分解:

$$X = X_1 \oplus X_2 \oplus \cdots \oplus X_n, \tag{3.2.17}$$

使得每个 X_i 是 T 的不变子空间(即 $TX_i \subset X_i, 1 \leqslant i \leqslant n$),且 $\sigma(T_i) = \sigma_i$,

$$Tx = \sum_i T_i x_i \quad \left(x = \sum x_i, x_i \in X_i\right), \tag{3.2.18}$$

此处 $T_i = T \mid X_i$ 看作 X_i 上的有界线性算子.

证 取充分小的 $\varepsilon > 0$,令

$$\Omega_i = \{\lambda \in \mathbf{C} : d(\lambda, \sigma_i) < \varepsilon\} \quad (1 \leqslant i \leqslant n),$$

使得 $\overline{\Omega}_i$ 互不相交(此种 ε 存在之理由基于 σ_i 的紧性,可参看定理 2.4.14 的证明).令 $\Omega = \bigcup_1^n \Omega_i$,则 Ω 是一开集,$\sigma(T) \subset \Omega$.以 $f_i(\lambda)$ 记 Ω_i 之特征函数,则 $f_i(\lambda) \in H(\Omega)$.令 $P_i = f_i(T)$,$X_i = P_i X$,今验证 $X_i (1 \leqslant i \leqslant n)$ 即合于定理要求.

(i) 验证式(3.2.17).显然有恒等式:

$$f_i(\lambda) f_j(\lambda) = \delta_{ij} f_i(\lambda), \sum_i f_i(\lambda) = 1 \quad (\lambda \in \Omega).$$

于是由定理 3.2.2 得出

$$P_i P_j = \delta_{ij} P_i, \quad \sum_i P_i = I \quad (1 \leqslant i, j \leqslant n). \tag{3.2.19}$$

特别,$P_i^2 = P_i (1 \leqslant i \leqslant n)$.由 $\sum P_i = I$ 得

$$X = \sum_i P_i X = \sum_i X_i,$$

这意味着每个 $x \in X$ 有分解

$$x = \sum x_i, \quad x_i \in X_i \quad (1 \leqslant i \leqslant n). \tag{3.2.20}$$

因 $X_i = P_i X$,故对式(3.2.20)中的 x_i 有 $y_i \in X$,使 $x_i = P_i y_i$,从而由式(3.2.19)

有

$$x_i = \sum_j \delta_{ij} P_i y_j = \sum_j P_i P_j y_j = P_i \sum_j x_j = P_i x.$$

这表明分解式(3.2.20)是惟一的,因而直和分解式(3.2.17)成立,且 P_i 就是从 X 到 X_i 的投影.因 $X_i = N\left(\sum_{j\neq i} P_j\right)(1 \leqslant i \leqslant n)$ 是 X 的闭子空间,故式(3.2.17)是拓扑直和.

(ii) 证式(3.2.18).任给 $x \in X$,令 $x_i = P_i x(1 \leqslant i \leqslant n)$,则

$$Tx = T\sum_i x_i = \sum_i Tx_i = \sum_i T_i x_i.$$

因 $TX_i = TP_i X = P_i TX \subset P_i X = X_i$,故 $X_i(1 \leqslant i \leqslant n)$ 是 T 的不变子空间.

(iii) 证 $\sigma(T_i) = \sigma_i(1 \leqslant i \leqslant n)$.只要证 $\sigma(T_1) = \sigma_1$.不妨设 $n = 2$(否则合并 $\sigma_2, \cdots, \sigma_n$).任取 $\lambda_0 \in \sigma_1$,令

$$\varphi(\lambda) = f_2(\lambda)/(\lambda_0 - \lambda),$$

则 $\varphi(\lambda) \in H(\Omega)$.于是由定理 3.2.2 有

$$P_2 = (\lambda_0 I - T)\varphi(T) = \varphi(T)(\lambda_0 I - T); \tag{3.2.21}$$

$$\varphi(T) = P_2\varphi(T), \quad \varphi(T)X \subset X_2. \tag{3.2.22}$$

式(3.2.22)表明 $\varphi(T) \mid X_2$ 可看作从 X_2 到 X_2 的算子,记作 B;由式(3.2.21)推出

$$I = P_2 \mid X_2 = (\lambda_0 I - T_2)B = B(\lambda_0 I - T_2),$$

可见 $\lambda_0 \overline{\in} \sigma(T_2)$.这推出 $\lambda_0 \in \sigma(T_1)$,否则,由 $\lambda_0 I - T_1$ 与 $\lambda_0 I - T_2$ 可逆将推出 $\lambda_0 I - T$ 可逆,这与 $\lambda_0 \in \sigma_1 \subset \sigma(T)$ 相矛盾.这就证得 $\sigma_1 \subset \sigma(T_1)$.

反之,若 $\lambda_0 \in \sigma(T_1)$,则必 $\lambda_0 \in \sigma(T)$(否则,$\lambda_0 I - T$ 可逆,这将推出 $\lambda_0 I - T_1$ 与 $\lambda_0 I - T_2$ 均可逆).若 $\lambda_0 \in \sigma_2$,则依上段所证将有 $\lambda \overline{\in} \sigma(T_1)$,得出矛盾.故必 $\lambda_0 \in \sigma_1$.因此 $\sigma(T_1) = \sigma_1$. □

当 $T_i \in L(X_i)(1 \leqslant i \leqslant n)$ 满足式(3.2.18)时约定

$$T = T_1 \oplus T_2 \oplus \cdots \oplus T_n. \tag{3.2.18'}$$

这样,空间的直和分解式(3.2.17)对应一个算子的"直和分解"式(3.2.18)′.定理 3.2.4 的意义在于,利用分解式(3.2.18)′,可将对算子 T 的研究归结为对算子 $T_i(1 \leqslant i \leqslant n)$ 的研究,而 T_i 可能是较简单或性质较好的算子.

一个简单的特例是:$\dim X = n$,$\sigma(T) = \{\lambda_1, \lambda_2, \cdots, \lambda_n\}$,$\lambda_i$ 互不相同.取 σ_i

$=\{\lambda_i\}(1\leqslant i\leqslant n)$ 应用定理 3.2.4 得到分解式(3.2.17)与(3.2.18)′,此时每个 X_i 必定是 X 的 1 维子空间,而 $\sigma(T_i)=\{\lambda_i\}$ 表明 T_i 是相似变换:

$$T_i x_i = \lambda_i x_i, x_i \in X_i \quad (1\leqslant i\leqslant n).$$

若 $0\neq e_i\in X_i$,则 $\{e_i\}$ 是 X 的基,T 关于基 $\{e_i\}$ 的矩阵正是对角形 $\mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$. 因此如上的谱分解称为**对角分解**.

至于在一般的 Banach 空间中应用定理 3.2.4,首先应取定一个如定理 3.2.4 所要求的分解 $\sigma(T)=\bigcup_1^n\sigma_i$. 这种分解可能根本不存在,当 $\sigma(T)$ 只含一点或 $\sigma(T)$ 是一连通集时就是如此. 在上述分解存在的情况下,也只有适当选定的分解才导致有意义的结果. 因此,毫无疑问,定理 3.2.4 的有效应用远不是一个简单的问题.

不过,我们还是可以举出并不十分复杂而又很有趣的例子,以下例子所包含的结论在动力系统理论中是基本的.

例 3.2.5(离散动力系统) 给定 $T\in L(X)$,由差分方程

$$x_0\in X;\quad x_{n+1}=Tx_n(n\geqslant 0) \tag{3.2.23}$$

给出 X 上的一个**离散动力系统**. 任给 $x_0\in X$,由式(3.2.23)决定一个序列 $\{x_n: n\geqslant 0\}$,称为系统从 x_0 发出的轨道. 基本的问题是,研究当 $n\to\infty$ 时轨道 $\{x_n\}$ 的渐近状态. 我们将发现,此问题的解答强烈地依赖于 T 的谱性质.

考虑如下的特殊情况:设 $\sigma(T)=\sigma_s\cup\sigma_u$,$\sigma_s$ 与 σ_u 均非空,且

$$\sigma_s=\{\lambda\in\sigma(T):0<|\lambda|<1\},$$

$$\sigma_u=\{\lambda\in\sigma(T):|\lambda|>1\}.$$

σ_s 与 σ_u 必为闭集(何故?),于是由定理 3.2.4 有分解:

$$X=X_s\oplus X_u; \tag{3.2.24}$$

$$T=T_s\oplus T_u, \tag{3.2.25}$$

其中 $T_s\in L(X_s)$,$T_u\in L(X_u)$,$\sigma(T_s)=\sigma_s$,$\sigma(T_u)=\sigma_u$. 因 $0\overline{\in}\sigma(T)$,故 T,T_s,T_u 均为可逆算子. 由 σ_s 与 σ_u 的定义,必有 $0<\rho<1$,使得

$$r_\sigma(T_s)<\rho,\quad r_\sigma(T_u^{-1})<\rho.$$

这结合谱半径公式(式(3.1.9))得出,当 n 充分大时有

$$\|T_s^n\|^{1/n}<\rho,\quad \|T_u^{-n}\|^{1/n}<\rho. \tag{3.2.26}$$

任给 $x\in X$,设 $x=x_s+x_u$,$x_s\in X_s$,$x_u\in X_u$. 由分解式(3.2.25)推出

$$T^n x=T^n x_s+T^n x_u=T_s^n x_s+T_u^n x_u.$$

今用式(3.2.26)推出以下结论:

(i) $T_s^n x_s \to 0; x_s \neq 0 \Rightarrow \| T_s^{-n} x_s \| \to \infty (n \to \infty)$. 事实上,

$$\| T_s^n x_s \| \leqslant \| T_s^n \| \| x_s \| \leqslant \rho^n \| x_s \| \to 0; \quad (\text{用式}(3.2.26))$$

因 $\| x_s \| = \| T_s^n T_s^{-n} x_s \| \leqslant \| T_s^n \| \| T_s^{-n} x_s \|$, 故当 $x_s \neq 0$ 时

$$\| T_s^{-n} x_s \| \geqslant \rho^{-n} \| x_s \| \to \infty (n \to \infty). \quad (\text{用式}(3.2.26))$$

(ii) 类似地, $T_u^{-n} x_u \to 0; x_u \neq 0 \Rightarrow \| T_u^n x_u \| \to \infty (n \to \infty)$.

综合以上结论,得出系统的轨道如图 3 2,为形象起见,图中用连续曲线代替了离散的轨道.

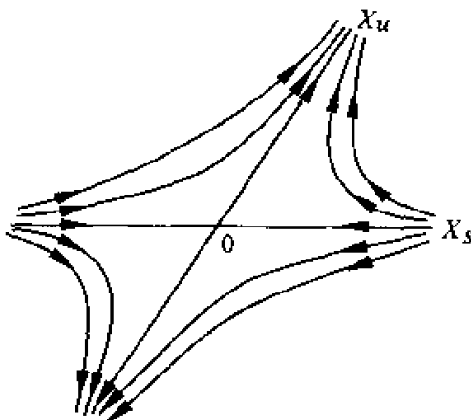


图 3-2

§3.3 紧线性算子

在本章开头,我们曾许诺在无限维空间中展开一个"特征值理论".现在让我们回顾一下,以上两节在多大程度上实现了这一目标.我们确实看到了一批深刻的结果,像谱半径公式、谱映射定理等,即使放在有限维空间中使用,也已超出线性代数的传统内容.这些结果的深刻性与普遍有效性值得高度肯定,但它们并不特别富有代数风格.尽管已建立的谱分解定理(定理 3.2.4)可看作是"对角标准形"的某种雏形,但它毕竟过于笼统,很难与线性代数中那些高度精细的结果进行对比.总之,我们还走得不远,这应归因于过分的一般化使得可利用的结构太少.这促使我们朝特殊的方向展开:研究紧线性算子与 Hilbert 空间上的线性算子,前者正是本节所要考虑的.

§3.3.1 紧线性算子概念

引进紧线性算子的一个基本考虑是,希望所研究的算子更加接近于有限维空

间中的线性算子. 设 X, Y 是两个赋范空间(本段中不必设 X 为复 Banach 空间), $T \in L(X, Y)$, 序列 $\{x_n\} \subset X$ 有界. 尽管 $\{Tx_n\}$ 有界,但当 $\dim Y = \infty$ 时并不能断定 $\{Tx_n\}$ 有收敛子列. 这种缺憾常常是处理无限维问题困难的根源. 一个自然的补救办法是,限定上述的 $\{Tx_n\}$ 必含收敛子列. 这就导向紧算子概念.

定义 3.3.1　设 $T: X \to Y$ 为线性算子. 若对任何有界序列 $\{x_n\} \subset X$, $\{Tx_n\}$ 有收敛子列,则称 T 为**紧线性算子**,或简称紧算子.

直接看出,线性算子 $T: X \to Y$ 是紧线性算子的充要条件是: T 映 X 中的有界集为 Y 中的相对紧集(参看定义 1.4.1),这个条件有时就用作紧线性算子的定义. 由此也得出,紧线性算子必定有界.

以 $\mathrm{CL}(X, Y)$ 记从 X 到 Y 的紧线性算子之全体. $\mathrm{CL}(X, X)$ 就简写作 $\mathrm{CL}(X)$. 显然,当 $\dim X < \infty$ 或 $\dim Y < \infty$ 时有 $\mathrm{CL}(X, Y) = L(X, Y)$. 因此,若 $T \in L(X, Y)$, $\dim R(T) < \infty$, 则 $T \in \mathrm{CL}(X, Y)$, 这样的 T 称为**有限秩算子**.

下面给出用弱收敛对紧线性算子的刻画.

定理 3.3.2　设 $T: X \to Y$ 是线性算子. 若 T 是紧算子,则它满足

$$x_n \rightharpoonup x \Rightarrow Tx_n \to Tx \quad (n \to \infty). \tag{3.3.1}$$

当 X 是自反空间时,式(3.3.1)是 T 为紧线性算子的充要条件.

证　设 T 是紧算子,在 X 中 $x_n \rightharpoonup x\ (n \to \infty)$. 若 $Tx_n \not\to Tx$, 则有 $\varepsilon > 0$ 与 $\{x_n\}$ 的子列 $\{x_{n_k}\}$, 使得

$$\| Tx_{n_k} - Tx \| \geqslant \varepsilon, \quad k = 1, 2, \cdots \tag{3.3.2}$$

因 $\{x_n\}$ 有界(用定理 2.5.2),故 $\{Tx_{n_k}\}$ 含收敛子列,不妨就设 $Tx_{n_k} \to y \in Y\ (k \to \infty)$. 另一方面,由 $x_{n_k} \rightharpoonup x$ 推出: $\forall v \in Y^*$, 有

$$\langle v, Tx_{n_k} \rangle = \langle T^* v, x_{n_k} \rangle \to \langle T^* v, x \rangle = \langle v, Tx \rangle,$$

可见 $Tx_{n_k} \rightharpoonup Tx$. 这推出 $y = Tx$, $Tx_{n_k} \to Tx$, 而这与条件(3.3.2)相矛盾. 可见 T 满足条件(3.3.1).

其次设 X 是自反空间, T 满足条件(3.3.1). 若 $\{x_n\} \subset X$ 是有界序列,则由定理 2.5.4 知 $\{x_n\}$ 含弱收敛子列. 设 $x_{n_k} \rightharpoonup x\ (k \to \infty)$, 则由条件(3.3.1)有 $Tx_{n_k} \to Tx$. 因此 T 是紧线性算子. □

紧线性算子有很好的运算性质,这概括于以下命题.

命题 3.3.3　(i) $\mathrm{CL}(X, Y)$ 是 $L(X, Y)$ 的子空间.

(ii) 紧线性算子与有界线性算子的乘积是紧线性算子,即若 $T \in \mathrm{CL}(X, Y)$, $A \in L(X, Z)$, $B \in L(W, X)$, 则

$$AT \in \mathrm{CL}(X,Z), \quad TB \in \mathrm{CL}(W,Y).$$

(iii) 若 Y 完备，则 $\mathrm{CL}(X,Y)$ 是 $L(X,Y)$ 的闭子空间.

(iv) 若 $T \in \mathrm{CL}(X,Y)$，则 $T^* \in \mathrm{CL}(Y^*,X^*)$.

证 (i)(ii) 的证明是直接的；(iv)的证明参看文献[10]§2.10，只证(iii).

设 $\{T_n\} \subset \mathrm{CL}(X,Y)$，$T \in L(X,Y)$，$\|T_n - T\| \to 0 (n \to \infty)$. 任给有界集 $A \subset X$，今证 TA 相对紧，为此只需证 TA 全有界(定理1.4.3). $\forall \varepsilon > 0$，取定 n，使 $\|T_n - T\| < \varepsilon$. 因 T_nA 全有界，故有有限集 $\{y_i\} \subset Y$，使得 $T_nA \subset \bigcup B_\varepsilon(y_i)$. 于是

$$TA \subset T_nA + B_\varepsilon(0) \subset \bigcup_i B_\varepsilon(y_i) + B_\varepsilon(0)$$
$$\subset \bigcup_i B_{2\varepsilon}(y_i),$$

这表明 TA 全有界，如所要证. □

简单地说，命题3.3.3表明：紧线性算子经线性运算、乘法运算、一致收敛的极限运算及取对偶之后，仍得紧线性算子. 这一结论对于紧线性算子的判别颇有用处. 下面就来看一些具体例子.

例3.3.4 (i) 设 $A = [a_{ij}]$ 是一无穷矩阵，$\sum_{i,j} |a_{ij}|^2 < \infty$，今指明 $A \in \mathrm{CL}(l^2)$（参看命题2.2.2(iii)）. 将 A 中前 n 行以外的元素换成零，所得之无穷矩阵记为 A_n，则对任给 $x \in l^2$，有

$$A_n x = (y_1, \cdots, y_n, 0, 0, \cdots),$$

故 A_n 是有限秩算子. 因 $A_n - A$ 的前 n 行均为零，故由命题2.2.2(iii)有

$$\|A_n - A\|_2^2 \leqslant \sum_{i>n} \sum_j |a_{ij}|^2 \to 0 \quad (n \to \infty).$$

于是由命题3.3.3(iii)得出 $A \in \mathrm{CL}(l^2)$.

(ii) 设 $J = [a,b] (a < b)$，积分算子 T 依式(2.2.8)，

$$\int_a^b \int_a^b |K(x,y)|^2 \mathrm{d}x\mathrm{d}y < \infty,$$

则 $T \in \mathrm{CL}(L^2(J))$. 与(i)对照，这一结论似乎可信，不过证明有所不同. 由定理3.3.2，只需证 T 满足条件(1). 设在 $L^2(J)$ 中 $u_n \rightharpoonup u$，今证 $\|Tu_n - Tu\|_2 \to 0$. 因

$$\|Tu_n - Tu\|_2^2 = \|Tu_n\|_2^2 + \|Tu\|_2^2 - \langle Tu_n, Tu \rangle - \langle Tu, Tu_n \rangle,$$

$$\langle Tu_n, Tu \rangle = \langle u_n, T^* Tu \rangle \to \langle u, T^* Tu \rangle = \|Tu\|_2^2 \quad (n \to \infty),$$

同理 $\langle Tu, Tu_n\rangle \to \|Tu\|_2^2$，故只要证 $\|Tu_n\|_2 \to \|Tu\|_2$，即

$$\lim_n \int_a^b |Tu_n(x)|^2 dx = \int_a^b |Tu(x)|^2 dx. \tag{3.3.3}$$

因

$$|Tu_n(x)|^2 = \left|\int_a^b K(x,y)u_n(y)dy\right|^2$$

$$\leqslant \|u_n\|_2^2 \int_a^b |K(x,y)|^2 dy \qquad (\text{用 Schwarz 不等式})$$

$$\leqslant \text{const} \int_a^b |K(x,y)|^2 dy \in L^1(J),$$

故可用控制收敛定理证式(3.3.3)，为此只要说明 $Tu_n \to Tu$.a.e.. 记 $K_x(y) = K(x,y)$，由 $K(x,y) \in L^2(J \times J)$ 知对几乎所有 $x \in J$ 有 $K_x \in L^2(J)$；对这样的 x 有

$$Tu_n(x) = \langle u_n, \overline{K}_x\rangle \to \langle u, \overline{K}_x\rangle = Tu(x),$$

这表明 $Tu_n \to Tu$，a.e.，如所要证.

将 J 换成任何 $\Omega \subset \mathbf{R}^n$，以上结论仍保持成立.

(iii) 设 J, T 如(ii)，但 $K(x,y)$ 连续，则 $T \in \mathrm{CL}(C(J))$. 为证此，只需指出：若 $A \subset C(J)$ 有界，则 TA 等度连续(从而依定理 1.4.6 知 TA 相对紧). 这由以下估计得出：

$$|Tu(x) - Tu(z)| = \left|\int_a^b [K(x,y) - K(z,y)]u(y)dy\right|$$

$$\leqslant \text{const} \int_a^b |K(x,y) - K(z,y)| dy \quad (\forall u \in A).$$

值得注意的是，若 X 是无限维 Banach 空间，则 X 上的单位算子 I 必非紧算子，否则，$\overline{B}_1(0) = I\overline{B}_1(0)$ 将为紧集(参看定理 1.4.8). 由此进而推出，任何 $T \in \mathrm{GL}(X)$ 必非紧算子(否则，由命题 3.3.3(ii)将推出 $I = TT^{-1}$ 是紧算子). 因此，$\mathrm{CL}(X)$ 作为 $L(X)$ 的闭的真子空间是疏集(参看第一章习题 14). 如此看来，从逻辑上说，在无限维空间中有界线性算子成为紧算子是极不寻常的. 不过，如例 3.3.4 所表明的，毕竟有一系列常用的线性算子是紧算子，这实在是一件可庆幸的事.

§3.3.2　特征值

重新假定 X 是复 Banach 空间. 如本节开头所预期的，X 上的紧线性算子非常

接近于有限维空间上的线性算子,这一事实在其"特征值理论"中得到最明显的体现.然而,整个理论的展开并不简单,在本段中仅给出一个基本的结果,它表明紧线性算子的谱有特别简单的结构.

任给 $T \in L(X)$,$\lambda \in \mathbf{C}$,约定 $T_\lambda = \lambda I - T$,$T_\lambda^* = \lambda I - T^* = (T_\lambda)^*$.注意 T_λ 与 T 总是可换的.

引理 3.3.5 设 $T \in CL(X)$,则 $R(T_1)$ 是 X 的闭子空间.

证 设 $\{x_n\} \subset X$,$T_1 x_n \to x (n \to \infty)$,今要证 $x \in R(T_1)$.若 $\{x_n\}$ 有界,则由 T 的紧性不妨设 $Tx_n \to y \in X (n \to \infty)$,于是

$$
\begin{aligned}
x &= \lim_n T_1(T_1 + T)x_n \qquad (\text{用 } T_1 + T = I)\\
&= T_1[\lim_n(T_1 x_n + Tx_n)]\\
&= T_1(x + y) \in R(T_1).
\end{aligned}
$$

因此下面只要考虑 $\{x_n\}$ 无界的情况.可设 $x_n \notin N(T_1)$(否则不妨设有 $\{x_{n_k}\} \subset N(T_1)$,由此将有 $x = 0 \in R(T_1)$),因此 $\rho_n \triangleq d(x_n, N(T_1)) > 0$.取 $y_n \in N(T_1)$,使得

$$
\rho_n \leqslant \| x_n - y_n \| \leqslant (1 + n^{-1})\rho_n \quad (n \in \mathbf{N}). \tag{3.3.4}
$$

若 $\{\rho_n\}$ 有界,则以 $x_n - y_n$ 取代 x_n,用前面所证得

$$
x = \lim_n T_1(x_n - y_n) \in R(T_1).
$$

若 $\{\rho_n\}$ 无界,则不妨设 $\rho_n \to \infty$(必要时过渡到一个子列).令

$$
z_n = (x_n - y_n)/\| x_n - y_n \|,
$$

则不妨设 $Tz_n \to z$(用 T 的紧性),因此

$$
\begin{aligned}
T_1 z &= \lim_n T_1 Tz_n = \lim_n TT_1 z_n \qquad (T, T_1 \text{ 可换})\\
&= \lim_n \| x_n - y_n \|^{-1} TT_1 x_n = 0,
\end{aligned}
$$

可见 $z \in N(T_1)$.注意 $z_n - (T_1 + T)z_n \to z (n \to \infty)$.另一方面,

$$
\begin{aligned}
\rho_n &\leqslant \| x_n - (y_n + \| x_n - y_n \| z) \| \qquad (y_n, z \in N(T_1))\\
&= \| x_n - y_n \| \| z_n - z \| \qquad (\text{用 } z_n \text{ 的定义})\\
&\leqslant \rho_n (1 + n^{-1}) \| z_n - z \|, \qquad (\text{用式}(3.3.4))
\end{aligned}
$$

这推出 $1 \leqslant \lim_n \| z_n - z \|$,得出矛盾.可见 $\{\rho_n\}$ 无界的情况不会出现.因此 $x \in$

$R(T_1)$ 得证. □

作了以上准备之后,现在来建立本段的主要结果.

定理 3.3.6　设 $T \in CL(X)$, 则 $\sigma(T)$ 是无非零聚点的可数集;若 $0 \neq \lambda \in \sigma(T)$, 则 $\lambda \in \sigma_p(T)$ 且 $\dim N(T_\lambda) < \infty$; 若 $\dim X = \infty$, 则 $0 \in \sigma(T)$.

定理证明不太简单,对之不感兴趣的读者不妨跳过它.

证　证明较长,分五步进行.

(i) 设 $0 \neq \lambda \in \sigma(T)$, 证 $\lambda \in \sigma_p(T)$, 这是证明的核心部分.不妨设 $\lambda = 1$ (否则以 $\lambda^{-1}T$ 代 T), 只要证 $N(T_1) \neq \{0\}$ (参看定义 3.1.2(iii)).用反证法:设 $N(T_1) = \{0\}$, 因而 $T_1 : X \to R(T_1)$ 是同构.因 $R(T_1)$ 作为 X 的闭子空间(用引理 3.3.5)是 Banach 空间,故由逆算子定理得出 $T_1^{-1} : R(T_1) \to X$ 有界.而由定理 2.4.13,这又推出 $R(T_1^*) = X^*$.

若能证 $N(T_1^*) = \{0\}$, 则有

$$R(T_1) = \overline{R(T_1)} = {}^{\perp}N(T_1^*) = X, \quad (\text{用式}(2.4.17))$$

从而 $T_1 : X \to X$ 为同构,这与假定 $1 \in \sigma(T)$ 相矛盾,因而 $1 \in \sigma_p(T)$ 得证.

于是只需证 $N(T_1^*) = \{0\}$. 仍用反证法:设 $N(T_1^*) \neq \{0\}$, 则有 $0 \neq u \in N(T_1^*)$. 由 $R(T_1^*) = X^*$, 可设 $u = T_1^* u_1$; 进而又可设 $u_1 = T_1^* u_2, \cdots$, 一般地, $u_n = T_1^* u_{n+1} (n \geqslant 1)$. 令 $N_n = N((T_1^*)^n)$, 则 N_n 是 X^* 的闭子空间,容易看出 $u_n \in N_{n+1} \setminus N_n$, 于是

$$N_n \subsetneqq N_{n+1} \quad (n \in \mathbf{N}).$$

由 Riesz 引理(引理 1.4.9),有 $v_n \in N_n$, 使得 $\| v_n \| = 1, d(v_n, N_{n-1}) \geqslant 1/2 (n \geqslant 2)$. 若 $n > m$, 则

$$(T_1^*)^{n-1}(T_1^* v_n + T^* v_m) = (T_1^*)^n v_n + T^*(T_1^*)^{n-1} v_m = 0,$$

故 $T_1^* v_n + T^* v_m \in N_{n-1}$. 于是

$$\| T^* v_n - T^* v_m \| = \| v_n - (T_1^* v_n + T^* v_m) \| \geqslant 1/2 \quad (n > m),$$

因而 $\{T^* v_n\}$ 不能含收敛子列,这与 T^* 为紧算子(用命题 3.3.3(iv))相矛盾.因此 $N(T^*) = \{0\}$, 如所要证.

(ii) 证 $\sigma(T)$ 无非零聚点.设 $\{\lambda_n\} \subset \sigma(T)$ 是一收敛序列,可设 λ_n 互不相同且 $\lambda_n \neq 0 (\forall n \in \mathbf{N})$. 令 $T_n = \lambda_n I - T$. 因 λ_n 是 T 的特征值,故有 $0 \neq x_n \in N(T_n) (n \in \mathbf{N})$, 即 x_n 是 T 关于 λ_n 的特征向量.如同在线性代数中一样,由 λ_n 互不相同可推出 $x_1, x_2, \cdots, x_n$ 线性无关.令 $X_n = \operatorname{span}\{x_1, x_2, \cdots, x_n\}$, 则如定理

1.4.8 的证明,由 Riesz 引理有 $y_n \in X_n$,使得 $\|y_n\| = 1, d(y_n, X_{n-1}) \geq 1/2 (n \geq 2)$. 令 $y_n = \sum_{i=1}^{n} \alpha_{ni} x_i$,则当 $n > m$ 时有

$$T_n y_n + T y_m = \sum_{i=1}^{n-1} \alpha_{ni}(\lambda_n - \lambda_i) x_i + \sum_{i=1}^{m} \alpha_{mi} \lambda_i x_i \in X_{n-1};$$

$$\begin{aligned} \| Ty_n - Ty_m \| &= \| \lambda_n y_n - (T_n y_n + T y_m) \| \\ &= |\lambda_n| \, \| y_n - \lambda_n^{-1}(T_n y_n + T y_m) \| \\ &\geq |\lambda_n| \, d(y_n, X_{n-1}) \geq |\lambda_n| / 2. \end{aligned}$$

因 $\{Ty_n\}$ 应含收敛子列,故只能有 $\lambda_n \to 0 (n \to \infty)$.

(iii) 证 $\sigma(T)$ 为可数集. 由 $\sigma(T)$ 无非零聚点推出,

$$S_n \triangleq \{\lambda \in \sigma(T) : 1/n \leq |\lambda| \leq n\} \quad (n \in \mathbf{N})$$

是有限集. 显然 $\sigma(T) \subset \{0\} \cup (\bigcup_1^\infty S_n)$, 故 $\sigma(T)$ 是可数集.

(iv) 设 $0 \neq \lambda \in \sigma(T)$, 证 $\dim N(T_\lambda) < \infty$. 为此,只要证 $N(T_\lambda)$ 中的单位球相对紧(参考定理 1.4.8),这由 T 为紧算子及以下事实推出:

$$\begin{aligned} &\{x \in N(T_\lambda) : \|x\| < 1\} \\ &= \{x \in B_1(0) : x = T(x/\lambda)\} \subset TB_{1/|\lambda|}(0). \end{aligned}$$

(v) 若 $\dim X = \infty$, 则 T 必不可逆,因此 $0 \in \sigma(T)$.

至此,定理证完. □

解释定理 3.3.6 的最简单例子是无穷对角矩阵

$$A = \operatorname{diag}(a_1, a_2, \cdots, a_i, \cdots),$$

其中 $a = (a_i) \in l^2$. 由例 3.3.4(i)知 $A \in CL(l^2)$. 类比于有限维的情况,自然猜想 $\sigma_p(A) = \{a_i\}$, 这实际上是对的: $\forall i \in \mathbf{N}$, 直接看出 $Ae_i = a_i e_i$ ($\{e_i\}$ 是 l^2 的标准基), 可见 $a_i \in \sigma_p(A)$, 且 e_i 就是A 关于a_i 的特征向量. 反之,若 $\lambda \in \sigma_p(A)$, $0 \neq x = (x_i) \in l^2, Ax = \lambda x$, 则必有某个 $i \in \mathbf{N}$ 使 $x_i \neq 0$, 因而从 $a_i x_i = \lambda x_i$ 得出 $\lambda = a_i$. 当不计重复时 $\{a_i\}$ 可以是有限集与无限可数集. 可见,定理 3.3.6 中的 $\sigma(T)$ 可以是无限可数集,此时它必以 $\lambda = 0$ 为其聚点.

§3.3.3 Riesz-Schauder 理论

作为一个无限维结果,定理 3.3.6 无疑是深刻而令人鼓舞的. 但它所对应的有

限维结论(读者试说说看)却平凡得不值一提.这就表明,推广通常特征值理论的工作还只是走出最初的一步.线性代数中那些更深入的结果,例如特征值的几何重数与代数重数、根子空间及矩阵的标准形等,在紧线性算子理论中将取什么形式,或者是否有其地位,都是有待回答且不容易的问题.在主要目的是介绍读者入门的本节中,我们不可能详尽地讨论这一切.但你肯定会对答案是什么感兴趣,这正是我们就要概括地给出的.按照传统说法,下面这个较长的结果称为 **Riesz-Schauder 理论**①.

定理 3.3.7 设 $T \in \mathrm{CL}(X), 0 \neq \lambda \in \mathbf{C}$.

(i) 若 $\lambda \in \sigma(T)$, 则存在 $r \geqslant 1$, 使得

$$N(T_\lambda) \subset N(T_\lambda^2) \subset \cdots \subset N(T_\lambda^r) = N(T_\lambda^k) \quad (\forall k \geqslant r); \tag{3.3.5}$$

$$R(T_\lambda) \supset R(T_\lambda^2) \supset \cdots \supset R(T_\lambda^r) = R(T_\lambda^k) \quad (\forall k \geqslant r); \tag{3.3.6}$$

$$X = N(T_\lambda^r) \oplus R(T_\lambda^r). \quad (\text{拓扑直和}) \tag{3.3.7}$$

式(3.3.5)与(3.3.6)中的包含均为真包含,且 $\dim N(T_\lambda^r) < \infty$.

(ii) 若 $\lambda \in \sigma(T)$, 则 $\dim N(T_\lambda) = \dim N(T_\lambda^*)$; 若 $\lambda \neq \mu$, 则

$$N(T_\lambda) \subset^\perp N(T_\mu^*).$$

(iii) 设 $y \in X, v \in X^*$, 对于互相联系的线性方程

$$\lambda x = Tx + y \qquad (3.3.8)_y$$

与

$$\lambda u = T^* u + v \qquad (3.3.9)_v$$

有以下结论:

方程 $(3.3.8)_y$ 有解 $x \in X \Leftrightarrow$ 对方程$(3.3.9)_0$ 的任何解 u 有 $u(y) = 0$;

方程 $(3.3.9)_v$ 有解 $u \in X^* \Leftrightarrow$ 对方程$(3.3.8)_0$ 的任何解 x 有 $v(x) = 0$;

对任何 $y \in X$ 方程 $(3.3.8)_y$ 恒有解$\Leftrightarrow$方程$(3.3.8)_0$ 仅有零解,此时 $x = T_\lambda^{-1} y$ 是方程 $(3.3.8)_y$ 的惟一解;

对任何 $v \in X^*$ 方程 $(3.3.9)_v$ 恒有解$\Leftrightarrow$方程$(3.3.9)_0$ 仅有零解,此时 $u = T_\lambda^{*-1} v$ 是方程 $(3.3.9)_v$ 的惟一解.

现在对上面这串颇长的结论略作解释.类比于矩阵理论,式(3.3.5)中的 $N(T_\lambda^r)$ 称为 T 关于定理特征值λ 的**根子空间**,它与特征子空间 $N(T_\lambda)$ 均为 T 的

---

① 通常所说的 Riesz-Schauder 理论,也包括定理 3.3.6.此处只是为了突出定理 3.3.6,将其预先单独处理了.

不变子空间,二者的维数分别相当于线性代数中的**代数重数**与**几何重数**.显然

$$\dim N(T_\lambda) \leqslant \dim N(T_\lambda^\nu),$$

这正好相当于熟知的线性代数结论:几何重数≤代数重数.

至于定理 3.3.7(iii),只要注意到方程 $(3.3.8)_y$ 有解 $\Leftrightarrow y \in R(T_\lambda)$,$x$ 是方程 $(3.3.8)_0$ 的解 $\Leftrightarrow x \in N(T_\lambda)$,就可将涉及方程 $(3.3.8)_y$ 与 $(3.3.9)_z$ 的四个结论缩写成:

$$R(T_\lambda) = {}^\perp N(T_\lambda^*); \tag{3.3.10}$$

$$R(T_\lambda^*) = N(T_\lambda)^\perp; \tag{3.3.11}$$

$$R(T_\lambda) = X \Leftrightarrow N(T_\lambda) = \{0\}; \tag{3.3.12}$$

$$R(T_\lambda^*) = X^* \Leftrightarrow N(T_\lambda^*) = \{0\}. \tag{3.3.13}$$

其中式(3.3.10)可由式(2.4.17)与引理 3.3.5 推出.仔细分析定理 3.3.6 之证的第(i)段,可以看出我们实际上已证明式(3.3.12)与式(3.3.13).至于等式(3.3.11)的证明,则需要某些更细致的分析.

等价关系(3.3.12)的意义在于,它将判定方程 $(3.3.8)_y$ 的可解性问题转化为判定对应的齐次方程只有零解的问题,后者往往要容易得多.试看一个简单例子.

例 3.3.8 考虑积分方程

$$\lambda u(x) = \int_0^1 u(y)\mathrm{d}y + v(x). \tag{3.3.14}$$

令

$$Tu = \int_0^1 u(y)\mathrm{d}y,$$

则 $T \in \mathrm{CL}(L^2(J))$ 且 $T \in \mathrm{CL}(C(J))$,$J = [0,1]$;方程 $(3.3.14)_v$ 可缩写成

$$\lambda u = Tu + v. \tag*{$(3.3.14)_0$}$$

设 $\lambda \neq 0$. 因 $Tu = \mathrm{const}$,故方程$(3.3.14)_0$ 只能有常数解 $u(x) \equiv a$. 以 $u(x) \equiv a$ 代入式$(3.3.14)_0$ 得 $\lambda a = a$. 若 $\lambda \neq 1$,则必 $a = 0$,因此方程$(3.3.14)_0$ 只有零解.于是由定理3.3.7(iii)得出结论:若 $\lambda \neq 0,1$,则对任何 $v \in L^2(J)$ (或 $v \in C(J)$),方程 $(3.3.14)_v$ 有惟一解 $u \in L^2(J)$ (或 $u \in C(J)$),这个解就是

$$u(x) = \lambda^{-1}[v(x) + b] \quad (x \in J),$$

其中 $b = \int_0^1 u(x)\mathrm{d}x$. 这与

$$\lambda\int_0^1 u(x)\mathrm{d}x = \int_0^1 v(x)\mathrm{d}x + b$$

一起解出 b，最终得到

$$u(x) = \frac{1}{\lambda}\left[v(x) + \frac{1}{\lambda - 1}\int_0^1 v(y)\mathrm{d}y\right] \quad (x \in J).$$

§3.4　Hilbert 空间上的有界线性算子

如所熟知，矩阵理论中最富有成果的一些内容，密切联系着内积、正交性等概念，这部分内容的无限维推广，自然涉及 Hilbert 空间上的线性算子. 实际上，线性算子谱论中最精致最优美的部分，正是关于 Hilbert 空间上的线性算子的. 本节只是介绍这一理论的最初等的部分.

以下设 H 是给定的复 Hilbert 空间.

§3.4.1　相伴算子

Hilbert 空间上的线性算子谱论之所以具有更丰富的内容，在很大程度上是因为有一个由内积定义的运算：

$$* : L(H) \to L(H), \quad T \to T^*. \tag{3.4.1}$$

这一新运算与线性运算、乘法一起，使得算子代数 $L(H)$ 具有很丰富的结构，这就为谱论的深入展开准备了条件.

定义 3.4.1　任给 $T \in L(H)$，由恒等式

$$\langle Tx, y\rangle = \langle x, T^* y\rangle \quad (x, y \in H) \tag{3.4.2}$$

决定一个算子 $T^* \in L(H)$，称它为 T 的**相伴算子**.

以上定义形式上固然很简单，但它有一个明显缺陷，即定义本身并未自动说明，式(3.4.2)确实决定一个算子 $T^* \in L(H)$，这就需要补充一番解释. 首先，对给定的 $y \in H$，由

$$u_y(x) = \langle Tx, y\rangle \quad (x \in H)$$

显然定义一个 $u_y \in H^*$，且 $\| u_y \| \leqslant \| T \| \, \| y \|$. 由定理 2.3.6，存在由 u_y（因而也就由 y）惟一决定的 $y^* \in H$，使得

$$u_y(x) = \langle x, y^* \rangle \quad (x \in H),$$

且 $\| y^* \| = \| u_y \|$. 只要令 $T^* y = y^*$，就得到一个确定的算子 $T^* : H \to H$，

$y \mapsto y^*$，它使得恒等式(3.4.2)满足，且

$$\| T^* y \| = \| u_y \| \leqslant \| T \| \, \| y \| .$$

余下只要说明 T^* 是线性的，即

$$T^*(\alpha y + \beta y_1) = \alpha T^* y + \beta T^* y_1 \quad (\alpha, \beta \in \mathbf{C}, y, y_1 \in H),$$

这由以下推演得出：

$$\begin{aligned} \langle x, T^*(\alpha y + \beta y_1) \rangle &= \langle Tx, \alpha y + \beta y_1 \rangle = \bar{\alpha} \langle Tx, y \rangle + \bar{\beta} \langle Tx, y_1 \rangle \\ &= \bar{\alpha} \langle x, T^* y \rangle + \bar{\beta} \langle x, T^* y_1 \rangle \qquad (\text{用(3.4.2)}) \\ &= \langle x, \alpha T^* y + \beta T^* y_1 \rangle \quad (\forall x \in H). \end{aligned}$$

这样，定义 3.4.1 的合法性得到证实.

T^* 这一记号大概会使你猜度：相伴算子是否就是对偶算子？形式上，恒等式(3.4.2)与式(2.3.16)确实高度类似. 容易说明，在 Hilbert 空间上，相伴算子与对偶算子有着很简单的联系，但二者在概念上并非完全一致. 考虑到对于 Hilbert 空间实际上只使用相伴算子，故记号 T^* 并不会与对偶算子相混淆. 而且，记住有关对偶算子的某些事实，对于展开相伴算子理论不无益处.

例 2.3.8 中关于对偶算子的两个例子略加改变，就成为相伴算子的两个典型例子. 首先设 $A = [a_{ij}] \in \mathbf{C}^{n \times n}$，则 $\forall x, y \in \mathbf{C}^n$，有

$$\begin{aligned} \langle Ax, y \rangle &= \sum_i \Big(\sum_j a_{ij} x_j \Big) \bar{y}_i = \sum_j x_j \sum_i a_{ij} \bar{y}_i \\ &= \sum_i x_i \overline{\sum_j \bar{a}_{ji} y_j} = \langle x, A^* y \rangle, \end{aligned}$$

这表明 $A^* = [\bar{a}_{ji}] = \bar{A}^{\mathrm{T}}$，即 A^* 原不过是 A 的**共轭转置**.

其次设 $T \in L(L^2(J))$ 是以 $K(x,y) \in L^2(J \times J)$ 为核的积分算子，$J = [a,b] (a < b)$ (参看命题 2.2.3(iii))，则 $\forall u, v \in L^2(J)$，有

$$\begin{aligned} \langle Tu, v \rangle &= \int_a^b \overline{v(x)} \mathrm{d}x \int_a^b K(x,y) u(y) \mathrm{d}y \\ &= \int_a^b u(y) \mathrm{d}y \int_a^b K(x,y) \overline{v(x)} \mathrm{d}x \qquad (\text{积分交换}) \\ &= \int_a^b u(x) \mathrm{d}x \overline{\int_a^b \overline{K(y,x)} v(y) \mathrm{d}y} \qquad (x, y \text{ 互换}) \\ &= \langle u, T^* v \rangle. \end{aligned}$$

由此可见，T^* 正是以 $\overline{K(y,x)}$ 为核的积分算子，而 $\overline{K(y,x)}$ 可看作 $K(x,y)$ 的“共轭转置”.

下面的命题综合了运算(1)的一些基本性质，它与命题 2.3.9 是高度类似的.

命题 3.4.2 对任给 $T, S \in L(H), \alpha, \beta \in \mathbf{C}$，成立以下等式：

$$(\alpha T + \beta S)^* = \bar{\alpha} T^* + \bar{\beta} S^*; \tag{3.4.3}$$

$$(TS)^* = S^* T^*, \quad (T^n)^* = (T^*)^n; \tag{3.4.4}$$

$$(T^{-1})^* = (T^*)^{-1} \quad (\text{若 } T \text{ 可逆}); \tag{3.4.5}$$

$$T^{**} = T; \tag{3.4.6}$$

$$\| T^* \| = \| T \|; \tag{3.4.7}$$

$$\| T \|^2 = \| TT^* \| = \| T^* T \|. \tag{3.4.8}$$

证 式(3.4.3)~(3.4.5)的验证是直接的(参照命题 2.3.9 中相应结论的证明). 其次，由

$$\begin{aligned} \langle x, T^{**} y \rangle &= \langle T^* x, y \rangle = \overline{\langle y, T^* x \rangle} \quad (\text{用(2)}) \\ &= \overline{\langle Ty, x \rangle} = \langle x, Ty \rangle \quad (\forall x, y \in H) \end{aligned}$$

得出 $T^{**} = T$. 为证(3.4.7)，首先指出(依定理 2.3.6 与式(2.3.1))

$$\| y \| = \sup_{\| x \| = 1} | \langle x, y \rangle | = \sup_{\| x \| = 1} | \langle y, x \rangle |. \tag{3.4.9}$$

于是

$$\begin{aligned} \| T^* \| &= \sup_{\| y \| = 1} \| T^* y \| \quad (\text{用式(2.1.7)}) \\ &= \sup_{\| y \| = 1} \sup_{\| x \| = 1} | \langle x, T^* y \rangle | \quad (\text{用式(3.4.9)}) \\ &= \sup_{\| x \| = 1} \sup_{\| y \| = 1} | \langle Tx, y \rangle | \quad (\text{用式(3.4.2)}) \\ &= \sup_{\| x \| = 1} \| Tx \| = \| T \|. \end{aligned}$$

对于式(3.4.8)，首先由

$$\begin{aligned} \| T^* T \| &\leqslant \| T^* \| \| T \| = \| T \|^2 \\ &= \sup_{\| x \| = 1} \| Tx \|^2 = \sup_{\| x \| = 1} \langle Tx, Tx \rangle \\ &= \sup_{\| x \| = 1} \langle x, T^* Tx \rangle \leqslant \| T^* T \| \end{aligned}$$

得 $\| T^* T \| = \| T \|^2$；其次有

$$\| TT^* \| = \| T^{**} T^* \| = \| T^* \|^2 = \| T \|^2, \quad (用式(3.4.7))$$

故式(3.4.8)得证. □

等式(3.4.3)~(3.4.6)在矩阵理论中是熟知的. 为理解式(3.4.7),式(3.4.8),可用一个更简单的比拟:若将 $a \in \mathbf{C}$ 等同于一阶矩阵 $[a]$, 则 $a^* = \bar{a}$. 这样,在 $L(\mathbf{C})(=\mathbf{C})$ 中, 等式(3.4.7),(3.4.8)成为熟知的:

$$|\bar{a}| = |a|; \quad |a\bar{a}| = |a|^2 \quad (a \in \mathbf{C}).$$

在一般情况下,亦可考虑将相伴算子比拟于共轭复数. 这一比拟看似过于简单化,但很具有启发性. 本节后面的内容,将充分说明这一点.

用稍形式化的语言说,命题 3.4.2 表明映射(3.4.1)是一个等距的共轭同构,它以自身为其逆映射(即 $T^{**} = T$). 这一事实的意义首先在于:一旦得出关于 $T, S \in L(H)$ 等的某个结论,就可能用命题 3.4.2 推出关于 T^*, S^* 等的相应结论. 例如,若已知 $\| T_n - T \| \to 0$, 则立得 $\| T_n^* - T^* \| \to 0 (n \to \infty)$; 若已知 $ST = TS$, 则必 $S^* T^* = T^* S^*$ (用(4))等.

§3.4.2 自伴算子与 U 算子

利用相伴算子概念,可界定几类重要的特殊算子,它们是 Hilbert 空间上算子理论的主要研究对象.

定义 3.4.3 设 $T \in L(H)$. 若 $TT^* = T^* T$, 则称 T 为**正规算子**;若 $T = T^*$, 则称 T 为**自伴算子**;若 $T^* = T^{-1}$, 则称 T 为 **U 算子**.

可将以上概念与线性代数中的下述约定对照:设 $A \in \mathbf{C}^{n \times n}$. 若 $AA^* = A^* A$, 则称 A 为**正规矩阵**;若 $A = A^*$, 则称 A 为 **Hermite 对称矩阵**(实的 Hermite 对称矩阵就是对称矩阵);若 $A^* = A^{-1}$, 则称 A 为 **U 矩阵**(实的 U 矩阵就是正交矩阵). 由此可见,定义 3.4.3 中的概念都有其有限维原型;从这些原型所获得的启发,对于展开自伴算子与 U 算子理论是十分有益的. 一个直接与矩阵类比而启示出的结论是:以 $K(x,y)$ 为核的积分算子(假定 $K(x,y)$ 平方可积)为自伴算子的充要条件是 $K(x,y) = \overline{K(y,x)}$. 一个更简单的比拟亦有用处:自伴算子可类比于实数 ($a = \bar{a} \Rightarrow a \in \mathbf{R}$), 而 U 算子则对应于模为 1 的复数($\bar{a} = a^{-1} \Rightarrow |a| = 1$). 这一比拟的好处在本节第 C 段中将更充分地显示出来.

从定义 3.4.3 直接看出,自伴算子与 U 算子都是正规算子. 不过,正规算子比自伴算子与 U 算子要宽泛得多,要得出关于正规算子的一个一般理论也就更不容易. 因此,我们主要考虑自伴算子与 U 算子. 另一方面,如下结论表明自伴算子处于更基本的地位:若 $A \in L(H)$ 是自伴算子,则 $T = e^{iA}$ 是 U 算子:

$$T^* = \left(\sum_{n=0}^{\infty} \frac{(iA)^n}{n!} \right)^*$$

$$= \sum_{n=0}^{\infty} \frac{(-\mathrm{i}A^*)^n}{n!} \qquad \text{(用式(3.4.3),(3.4.4),(3.4.7))}$$

$$= \sum_{n=0}^{\infty} \frac{(-\mathrm{i}A)^n}{n!} = \mathrm{e}^{-\mathrm{i}A} = T^{-1}.$$

更细致的分析可证明,若 $T \in L(H)$ 是 U 算子,且 $\sigma(T)$ 不充满复平面上的单位圆周,则必有自伴算子 $A \in L(H)$,使得 $T = \mathrm{e}^{\mathrm{i}A}$. 这就可能将自伴算子的有关结论直接转移到 U 算子.有趣的是,以上结论可以从 $|\mathrm{e}^{\mathrm{i}a}| = 1 (a \in \mathbf{R})$ 这一简单事实得到启发.

命题 3.4.4　设 $T \in L(H)$,则以下结论成立:

(i) T 是正规算子 $\Leftrightarrow \| Tx \| = \| T^* x \| (\forall x \in H)$.

(ii) T 是自伴算子 $\Leftrightarrow \langle Tx, x\rangle \in \mathbf{R} (\forall x \in H)$.

(iii) T 是 U 算子 $\Leftrightarrow T : H \to H$ 是等距同构.

证　首先,用式(1.5.10)建立恒等式:

$$4\langle Tx, y\rangle = \langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle + \mathrm{i}\langle T(x+\mathrm{i}y), x+\mathrm{i}y\rangle - \mathrm{i}\langle T(x-\mathrm{i}y), x-\mathrm{i}y\rangle. \tag{3.4.10}$$

由式(3.4.10)推出:若 $\langle Tx, x\rangle \equiv 0 (x \in H)$,则 $\langle Tx, y\rangle \equiv 0 (x, y \in H)$,从而 $Tx \equiv 0$,即 $T = 0$. 由此又推出:若 $\langle Tx, x\rangle \equiv \langle Sx, x\rangle (x \in H)$,则 $T = S$. 这表明 T 完全由二次泛函 $x \to \langle Tx, x\rangle (x \in H)$ 所决定.

利用以上结论,有

$$TT^* = T^*T \Leftrightarrow \langle TT^*x, x\rangle \equiv \langle T^*Tx, x\rangle$$

$$\Leftrightarrow \langle T^*x, T^*x\rangle \equiv \langle Tx, Tx\rangle, \quad \text{(用式(3.4.2))}$$

这得出结论(i).其次,

$$T = T^* \Leftrightarrow \langle Tx, x\rangle \equiv \langle T^*x, x\rangle$$

$$\Leftrightarrow \langle Tx, x\rangle \equiv \overline{\langle Tx, x\rangle}, \quad \text{(用式(3.4.2))}$$

即 $T = T^* \Leftrightarrow \langle Tx, x\rangle \in \mathbf{R} (\forall x \in H)$,这得出结论(ii).类似地,

$$T^* = T^{-1} \Leftrightarrow TT^* = T^*T = I$$

$$\Leftrightarrow \| Tx \| \equiv \| T^*x \| \equiv \| x \|$$

$$\Leftrightarrow T \text{ 是等距同构},$$

这得出结论(iii).　□

现在考虑本节的中心课题：自伴算子与 U 算子的谱. 由线性代数知道，Hermite 对称矩阵有实特征值；U 矩阵的特征值有绝对值 1. 这些结论可推广为更一般的定理.

定理 3.4.5 设 $T \in L(H)$.

(i) 若 T 是正规算子，则 $r_\sigma(T) = \|T\|$.

(ii) 若 T 是自伴算子，则 $\sigma(T) \subset \mathbf{R}$；设 $[m, M]$ 是包含 $\sigma(T)$ 的最小闭区间，则

$$m = \inf_{\|x\|=1} \langle Tx, x \rangle, \quad M = \sup_{\|x\|=1} \langle Tx, x \rangle; \tag{3.4.11}$$

$$\|T\| = \max\{|m|, |M|\} = \sup_{\|x\|=1} |\langle Tx, x \rangle|. \tag{3.4.12}$$

(iii) 若 T 是 U 算子，则 $\forall \lambda \in \sigma(T)$，有 $|\lambda| = 1$，即 $\sigma(T) \subset S^1$，S^1 记单位圆周.

证 (i) 由 $TT^* = T^*T$ 及式(3.4.8)有

$$\begin{aligned} \|T\|^4 = \|T^*T\|^2 &= \|T^*T(T^*T)^*\| \\ &= \|T^2(T^2)^*\| = \|T^2\|^2, \end{aligned}$$

故得 $\|T^2\| = \|T\|^2$. 进而归纳地得出 $\|T^{2^n}\| = \|T\|^{2^n} (n \geqslant 1)$，于是

$$r_\sigma(T) = \lim_n \|T^{2^n}\|^{1/2^n} = \|T\|. \quad (\text{用谱半径公式})$$

(ii) 首先证 $\sigma(T) \subset \mathbf{R}$. 设 $\lambda = \alpha + \mathrm{i}\beta \in \sigma(T)$，$\alpha, \beta \in \mathbf{R}$，今证 $\beta = 0$. $\forall t \in \mathbf{R}$，有 $\lambda + \mathrm{i}t \in \sigma(T + \mathrm{i}tI)$（用定理 3.2.3），于是

$$\begin{aligned} \alpha^2 + (\beta + t)^2 &\leqslant \|T + \mathrm{i}tI\|^2 && (\text{用式(3.1.13)}) \\ &= \|(T + \mathrm{i}tI)(T + \mathrm{i}tI)^*\| && (\text{用式(3.4.8)}) \\ &= \|T^2 + t^2 I\| \leqslant \|T\|^2 + t^2, && (\text{用 } T = T^*) \end{aligned}$$

这推出 $2\beta t \leqslant \|T\|^2 - \alpha^2 - \beta^2$. 要使此不等式对任何 $t \in \mathbf{R}$ 成立，必须 $\beta = 0$.

其次证 $\sigma(T) \subset [m, M]$，m, M 依式(3.4.11). 由式(3.4.11)推出：

$$m\|x\|^2 \leqslant \langle Tx, x \rangle \leqslant M\|x\|^2 \quad (x \in H). \tag{3.4.13}$$

只需证 $\lambda > M \Rightarrow \lambda \in \rho(T)$（同理将有 $\lambda < m \Rightarrow \lambda \in \rho(T)$）. 设 $\lambda > M$，令 $T_\lambda = \lambda I - T$，则 $\forall x \in H$，有

$$\begin{aligned} \|T_\lambda x\| \|x\| &\geqslant \langle T_\lambda x, x \rangle = \lambda \|x\|^2 - \langle Tx, x \rangle \\ &\geqslant (\lambda - M)\|x\|^2, \quad (\text{用式(3.4.13)}) \end{aligned}$$

于是 $\|T_\lambda x\| \geqslant (\lambda - M)\|x\|$. 因 $\lambda - M > 0$，故由命题 2.1.8 推出 $T_\lambda : X \to$

$R(T_\lambda)$ 是拓扑同构,因此 $R(T_\lambda)$ 必为 H 的闭子空间. 为证 $\lambda \in \rho(T)$, 只要证 $R(T_\lambda) = H$; 为此又只要证 $R(T_\lambda)^\perp = \{0\}$. 若 $y \in R(T_\lambda)^\perp$, 则 $\forall x \in H$, 有

$$0 = \langle T_\lambda x, y\rangle = \langle \lambda x, y\rangle - \langle Tx, y\rangle$$

$$= \langle x, \lambda y - Ty\rangle, \qquad (\text{用 } T = T^*)$$

因而 $Ty = \lambda y, y \in N(T_\lambda) = \{0\}$. 故 $R(T_\lambda)^\perp = \{0\}$, 如所要证.

然后证 $\|T\| = \beta$, 此处 β 记式(3.4.12)的右端. 显然 $\beta = \max\{|m|, |M|\}$, 且

$$|\langle Tx, x\rangle| \leqslant \beta\|x\|^2 \quad (x \in H). \tag{3.4.14}$$

由式(3.4.12)直接看出 $\beta \leqslant \|T\|$, 故只需证 $\beta \geqslant \|T\|$. 因

$$\|T\| = \sup_{\|x\|=1}\|Tx\| = \sup_{\|x\|=\|y\|=1}|\langle Tx, y\rangle|,$$

故只要对取定的 $x, y \in H, \|x\| = \|y\| = 1$, 证 $|\langle Tx, y\rangle| \leqslant \beta$. 令 $\varepsilon = \operatorname{sgn}\langle Tx, y\rangle$ ①,则

$$\begin{aligned}
4|\langle Tx, y\rangle| &= 4\langle Tx, \bar{\varepsilon}y\rangle = 4\mathrm{Re}\langle Tx, z\rangle && (z = \bar{\varepsilon}y)\\
&= \langle T(x+z), x+z\rangle - \langle T(x-z), x-z\rangle && (\text{用式}(3.4.10))\\
&\leqslant \beta(\|x+z\|^2 + \|x-z\|^2) && (\text{用式}(3.4.14))\\
&= 2\beta(\|x\|^2 + \|z\|^2) \leqslant 4\beta, && (\text{用中线公式})
\end{aligned}$$

故得 $|\langle Tx, y\rangle| \leqslant \beta$, 如所要证.

最后证 $M \in \sigma(T)$ (同理将有 $m \in \sigma(T)$). 不妨设 $0 \leqslant m \leqslant M$ (否则以 $T - mI$ 代 T),则由式(3.4.12)与已证的(i)有 $M = \|T\| = r_\sigma(T)$. 因 $\sigma(T)$ 是闭集,故必 $M \in \sigma(T)$. 于是(ii)得证.

(iii) 首先,由 $TT^* = I$ 及式(3.4.8)有

$$1 = \|TT^*\| = \|T\|^2,$$

故 $\|T\| = \|T^*\| = 1$ (用式(3.4.7)). 任给 $\lambda \in \sigma(T)$, 有 $|\lambda| \leqslant \|T\| = 1$; 另一方面,从 $\lambda^{-1} \in \sigma(T^{-1}) = \sigma(T^*)$ 得 $|\lambda^{-1}| \leqslant \|T^*\| = 1$, 因此 $|\lambda| = 1$, 如所要证. □

定理 3.4.5 中最重要的结论是(ii),今用一些具体例子来加以说明.

例 3.4.6 (i) 设 $A = A^* \in \mathbf{C}^{n\times n}, \sigma(A) = \{\lambda_1, \lambda_2, \cdots, \lambda_n\}, \lambda_1 \leqslant \lambda_2 \leqslant \cdots$

---

① 参看第 77 页注①。

$\leqslant \lambda_n$. 由线性代数熟知有 U 矩阵 Q, 使得

$$Q^* AQ = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n).$$

注意到 U 变换 $x = Qy$ 是等距同构(3.4.4(iii)),得出

$$\min_{|x|=1}\langle Ax, x\rangle = \min_{|y|=1}\langle AQy, Qy\rangle = \min_{|y|=1} y^* Q^* AQy$$

$$= \min_{|y|=1}\sum_i \lambda_i \mid y_i \mid^2 = \lambda_1 \quad (y^* = \bar{y}^{\mathrm{T}});$$

类似地, $\max\limits_{|x|=1}\langle Ax, x\rangle = \lambda_n$. 这与定理 3.4.5(ii)的结论正好相合.

由此可见,Hermite 对称矩阵的特征值性质,为定理 3.4.5(ii)提供了有限维原型.至于在无限维空间中,自伴算子 T 的谱则可能是 $[m, M]$(m, M 依式(3.4.11))的任何闭子集(但保持 $m, M \in \sigma(T)$);特别,可能有 $\sigma(T) = [m, M]$. 这可通过下面的例子来说明.

(ii) 设 $J = [a, b]$($a < b$), $\varphi \in L^\infty(J)$. 定义

$$Tu = \varphi u, \quad (u \in L^2(J)) \tag{3.4.15}$$

则显然 $T \in L(L^2(J))$, 且 $\| T \| \leqslant \| \varphi \|_\infty$. $\forall u, v \in L^2(J)$, 有

$$\langle Tu, v\rangle = \langle \varphi u, v\rangle = \langle u, \bar{\varphi} v\rangle = \langle u, T^* v\rangle,$$

可见

$$T^* u = \bar{\varphi} u \quad (u \in L^2(J)). \tag{3.4.16}$$

这就表明: T 是自伴算子 $\Leftrightarrow \varphi$ 是实函数, T 是 U 算子 $\Leftrightarrow \bar{\varphi} = \varphi^{-1} \Leftrightarrow |\varphi| = 1$. 若 φ 是实连续函数,则可验证二次泛函

$$\langle Tu, u\rangle = \int_a^b \varphi(x) \mid u(x) \mid^2 \mathrm{d}x$$

在单位球面 $\| u \|_2 = 1$ 上的上、下确界分别为 φ 在 J 上的最大值 M 与最小值 m, $[m, M]$ 就是包含 $\sigma(T)$ 的最小闭区间.

为确定起见,下面设 φ 为线性函数, $\varphi(a) = m, \varphi(b) = M, m < M$. 今指明 $\sigma(T) = [m, M]$. 事实上, $\forall \lambda \in \mathbf{R}, v \in L^2(J)$, 从方程

$$\lambda u - \varphi u = v$$

可惟一地解出

$$u = (\lambda - \varphi)^{-1} v.$$

若 $\lambda \,\overline{\in}\, [m, T]$, 则上述的 $u \in L^2(J)$, 可见 $\lambda I - T$ 可逆,因而 $\lambda \in \rho(T)$. 若 λ

$\in [m, T]$，则取 $v \equiv 1$ 有

$$\int_a^b |u(x)|^2 \mathrm{d}x = \int_a^b \frac{\mathrm{d}x}{|\lambda - \varphi(x)|^2}$$

$$= \int_m^M \frac{\mathrm{d}\varphi^{-1}(y)}{(\lambda - y)^2} = \infty,$$

即 $u \overline{\in} L^2(J)$，可见 $\lambda \in \sigma(T)$.

另一方面，若 $|\varphi| = 1$，则当 $\lambda \in \mathbf{C}$，$|\lambda| \neq 1$ 时，对任给 $v \in L^2(J)$ 有

$$(\lambda - \varphi)^{-1} v \in L^2(J),$$

因而 $\lambda \in \rho(T)$. 这表明 $\sigma(T) \subset S^1$. 取 $J = [0, 2\pi]$，$\varphi(x) = \mathrm{e}^{\mathrm{i}x}$ 说明，对于一个 U 算子 T 可能有 $\sigma(T) = S^1$.

形如式(3.4.15)的算子可看作有限维对角形的推广，它用来解释 Hilbert 空间中的线性算子概念是很有效的，后面还要用到类似的例子.

§3.4.3 正算子

在线性代数中，对称矩阵的正定性是一重要概念，下面来推广此概念. 如同在线性代数中一样，对于自伴算子 $T \in L(H)$，总是将它与二次泛函 $\langle Tx, x\rangle$ 联系起来考虑. 显然，二次泛函正是线性代数中二次型的推广：若 $A = A^{\mathrm{T}} \in \mathbf{R}^{n\times n}$，则

$$\langle Ax, x\rangle = x^{\mathrm{T}} Ax \quad (x \in \mathbf{R}^n)$$

就是以 A 为矩阵的二次型；若 $A = A^* \in \mathbf{C}^{n\times n}$，则 $\langle Ax, x\rangle (x \in \mathbf{C}^n)$ 就是以 A 为矩阵的 Hermite 二次型.

定义 3.4.7 设 $T, S \in L(H)$ 是自伴算子. 若 $\langle Tx, x\rangle \geqslant 0 (\forall x \in H)$，则称 T 为**正算子**，记作 $T \geqslant 0$；若 $T - S \geqslant 0$，则约定 $T \geqslant S$ 或 $S \leqslant T$.

设 m 依式(3.4.11)，则 $\langle Tx, x\rangle \geqslant 0 (\forall x \in H) \Leftrightarrow m \geqslant 0 \Leftrightarrow \sigma(T) \subset [0, \infty)$（参看定理3.4.5(ii)）. 可见，自伴算子 T 是正算子的充要条件是其谱值非负. 若 $m > 0$，则 $\langle Tx, x\rangle \geqslant 0$ 提升为更强的不等式（依式(3.4.13)）：

$$\langle Tx, x\rangle \geqslant m \| x \|^2 \quad (x \in H),$$

此时不妨称 T 为**正定算子**. 这样，自伴算子 T 正定 $\Leftrightarrow \sigma(T) \subset (0, \infty)$，这与线性代数中的结果一致. 而正算子则与线性代数中的正半定矩阵相当.

容易验证，自伴算子相加及与实数相乘仍得自伴算子. 这就表明，H 上的自伴算子之全体构成一**实**向量空间，记作 $L_s(H)$. 在很多方面，$L_s(H)$ 与 $L(H)$ 的关系，犹如 $\mathbf{R}$ 与 $\mathbf{C}$ 的关系. 一个简单而有趣的事实是每个 $T \in L(H)$ 有惟一分解：

$$T = A + \mathrm{i}B, \quad A, B \in L_s(H)$$

且

$$A = (T + T^*)/2, \quad B = (T - T^*)/2\mathrm{i},$$

因而 $T^* = A - \mathrm{i}B$. 你看出分解式 $T = A + \mathrm{i}B$ 与复数的表示 $z = a + \mathrm{i}b$ 正好相当. 有些作者甚至就称 A, B 为 T 的"实部"与"虚部". 基于这一类比,你更容易理解定义 3.4.7 中所定义的**序关系**$\leqslant$,它正好与实数之间的序关系相当,虽然不尽相同. 借助于序$\leqslant$,可以在 $L_s(H)$ 中建立丰富的结构,从而推广实域 $\mathbf{R}$ 上的许多熟知事实. 可惜,我们已没有太多篇幅来详尽讨论这一切了,下面只是列举一些最基本的结论,以使读者有一个大概印象.

定理 3.4.8 对于 H 上的自伴算子,以下结论成立:

(i) $\leqslant$是一个半序,这意味着 $T \leqslant T$ (自反性), $R \leqslant S \leqslant T \Rightarrow R \leqslant T$ (可传性), $T \leqslant S \leqslant T \Rightarrow S = T$ (反对称性).

(ii) $\leqslant$是线性的,这意味着 $T \leqslant S \Rightarrow \alpha T \leqslant \alpha S (\alpha \geqslant 0)$, $T_i \leqslant S_i (i = 1, 2) \Rightarrow T_1 + T_2 \leqslant S_1 + S_2$.

(iii) $\leqslant$是可乘的,这意味着,若 $A \geqslant 0, T \geqslant S$, A 与 T, S 可换,则 $AT \geqslant AS$; 特别,若 $A \geqslant 0, T \geqslant 0$, A 与 T 可换,则 $AT \geqslant 0$.

(iv) $\leqslant$是连续的,这意味着 $T_n \leqslant S_n (\forall n \in \mathbf{N}) \Rightarrow \lim\limits_n T_n \leqslant \lim\limits_n S_n$, 此处极限是在强收敛意义下取的,且假定其存在.

(v) 单调有界收敛原理. 若 $T_n \leqslant T_{n+1} \leqslant A (\forall n \in \mathbf{N})$, 则 $\{T_n\}$ 强收敛于某个 $T \in L_s(H)$, 且 $T \leqslant A$. 将$\leqslant$改为$\geqslant$,所述结论保持成立.

(vi) 序区间. $0 \leqslant T \leqslant I \Leftrightarrow \sigma(T) \subset [0,1]$; $\forall \alpha > 0$, 有

$$-\alpha I \leqslant T \leqslant \alpha I \Leftrightarrow \sigma(T) \subset [-\alpha, \alpha] \Leftrightarrow \| T \| \leqslant \alpha;$$

$$T \leqslant T_n \leqslant S \Rightarrow \sup_n \| T_n \| < \infty.$$

最后这个性质意味着:对于自伴算子序列,序有界推出范数有界.

(vii) 正平方根. 若 $T \geqslant 0$, 则存在惟一的 $A \geqslant 0$, 使得 $A^2 = T$; 这样的 A 称为 T 的**正平方根**,记作 $T^{1/2}$; $TS = ST \Rightarrow T^{1/2}S = ST^{1/2}$.

以上结论的证明并不都是平凡的(例如参考文献[10] §3.6),但结论本身则简单明白,类比于实数更可赋予某种直观性.

若 m, M 依(11),则 $mI \leqslant T \leqslant MI$. 因此, M 与 m 也分别称为 T 的上界与下界.

任给 $T \in L_s(H)$, 显然 $T^2 \geqslant 0$, 约定 $|T| = (T^2)^{1/2}$,

$$T^+ = \frac{1}{2}(|T| + T), \quad T^- = \frac{1}{2}(|T| - T).$$

我们自然联想到关于实函数的类似记号,因而不免将 $|T|, T^+$, 与 T^- 分别认作 T 的"绝对值"、"正部"与"负部". 当然,这只是一种比拟,但这一比拟使得与之有关的结论更好理解. 关于 $|T|, T^+, T^-$ 的以下性质容易从其定义直接推出.

命题 3.4.9　设 $T \in L_s(H)$, 则以下结论成立:

(i) $T = T^+ - T^-$, $|T| = T^+ + T^-$.

(ii) 若 $T \geqslant 0$, 则 $|T| = T^+ = T, T^- = 0$; 若 $T \leqslant 0$, 则 $|T| = T^- = -T$, $T^+ = 0$.

(iii) $|-T| = |T|, (-T)^+ = T^-, (-T)^- = T^+$.

(iv) $T^+T^- = T^-T^+ = 0$.

(v) 若 T 与 $S \in L(X)$ 可换,则 $|T|, T^+, T^-$ 均与 S 可换.

利用由(15)定义的算子 T, 可得到以上结果的最清晰的解释. 设 $\varphi \in L^\infty(J)$ 是实函数, $Tu = \varphi u (u \in L^2(J))$, 则 $|T|u = |\varphi| u, T^+ u = \varphi^+ u, T^- u = \varphi^- u$. 可见 $|T|, T^+$ 与 T^- 分别对应 φ 在通常意义下的绝对值、正部与负部. 在这种情况下,命题 3.4.8 的结论完全是平凡的.

§3.4.4　正投影

首先回忆一下,若 $H = A \oplus A^\perp$ 是正交分解,则由此分解决定的从 H 到 A 的投影称为正投影算子或正投影,通常记作 P_A; $\forall x \in H, P_A x$ 就是 x 在 A 中的最佳逼近(参看定理 1.5.10). 现在系统地研究正投影之间的运算与相互关系. 为此,首先给出正投影的以下刻画.

命题 3.4.10　$T \in L(H)$ 是正投影 $\Leftrightarrow T^2 = T = T^*$.

证　首先设 $T = P_A$, A 是 H 的闭子空间. 直接看出 $T^2 = T$. 其次,由

$$\begin{aligned}\langle Tx, y\rangle &= \langle Tx, y - Ty\rangle + \langle Tx, Ty\rangle \\ &= \langle Tx, Ty\rangle \qquad (y - Ty \in A^\perp) \\ &= \langle Tx - x, Ty\rangle + \langle x, Ty\rangle \\ &= \langle x, Ty\rangle \qquad (\forall x, y \in H)\end{aligned}$$

推出 $T = T^*$. 反之,设 $T^2 = T = T^*$, 令 $A = R(T), B = R(I - T)$, 则直接看出 $H = A + B$. $\forall x, y \in H$, 有

$$\begin{aligned}\langle Tx, y - Ty\rangle &= \langle Tx, y\rangle - \langle Tx, Ty\rangle \\ &= \langle Tx, y\rangle - \langle Tx, y\rangle = 0, \quad (\text{用 } T^2 = T = T^*)\end{aligned}$$

这表明 $B \subset A^\perp$. 因 $A = N(I - T)$ 是闭子空间,故有 $H = A \oplus A^\perp$ (依定理 1.5.10),因而 $B = A^\perp$, $T = P_A$ 是正投影. □

设 $P \in L(H)$ 是正投影,现在利用命题 3.4.10 来推出 P 的某些性质. 由 $P^2 = P$ 与定理 3.2.3 推出: $\forall \lambda \in \sigma(P)$, 有 $\lambda^2 - \lambda = 0$, 因此 $\sigma(P) \subset \{0,1\}$. 若 $\sigma(P) = \{0\}$, 则 $\|P\| = 0$ (用(12)),从而 $P = 0$; 若 $\sigma(P) = \{1\}$, 则 P 可逆,因而必 $P = I$; 若 $P \neq 0, I$, 则 $\sigma(P) = \{0,1\}$, $\|P\| = 1$ (用(12)). 在任一种情况下有

$$0 \leqslant \langle Px, Px \rangle = \langle Px, x \rangle \leqslant \langle x, x \rangle,$$

这表明 $0 \leqslant P \leqslant I$. 说明正投影的最直观的例子是对角形

$$P = \mathrm{diag}(a_1, a_2, \cdots, a_n), a_i = 0 \text{ 或 } 1 (1 \leqslant i \leqslant n),$$

它显然满足 $P^2 = P = P^*$.

在一定条件下,正投影的积、和或差仍为正投影. 准确地说,就是以下命题.

命题 3.4.11 设 $P_i \in L(H)$ 是正投影, $A_i = R(P_i)(i = 1,2)$, 则以下结论成立:

(i) P_1P_2 是正投影 $\Leftrightarrow P_1P_2 = P_2P_1 \Rightarrow R(P_1P_2) = A_1 \cap A_2$.

(ii) $P_1 + P_2$ 是正投影 $\Leftrightarrow P_1P_2 = 0 \Leftrightarrow A_1 \perp A_2 \Rightarrow R(P_1 + P_2) = A_1 + A_2$.

(iii) $P_2 - P_1$ 是正投影 $\Leftrightarrow P_1P_2 = P_2P_1 = P_1 \Leftrightarrow A_1 \subset A_2 \Leftrightarrow P_1 \leqslant P_2$.

以上结论的证明并不十分复杂. 但我们宁可放过这些证明,以便尽快看到本节的最终结果:自伴算子的谱分解.

定义 3.4.12 若正投影族 $\{E_\lambda : \lambda \in \mathbf{R}\} \subset L(H)$ 满足条件:

(i) $\lambda \leqslant \mu \Rightarrow E_\lambda \leqslant E_\mu$;

(ii) 当 $\lambda < \alpha$ 时 $E_\lambda = 0$, 当 $\lambda \geqslant \beta$ 时 $E_\lambda = I$;

(iii) 当 $\mu \downarrow \lambda$ 时 $E_\mu x \to E_\lambda x (\forall x \in H)$.

则称 $\{E_\lambda\}$ 为区间 $[\alpha, \beta]$ 上的**谱族**.

为表述谱分解定理,还需要关于算子值函数的 RS 积分(即 Riemann-Stieltjes 积分)概念. 设 $f(\lambda)$ 与 $E(\lambda)$ 分别为区间 $[\alpha,\beta]$ 上的数值函数与算子值函数,则定义 $f(\lambda)$ 关于 $E(\lambda)$ 的 RS 积分为

$$\int_\alpha^\beta f(\lambda) \mathrm{d}E(\lambda) = \lim_{\max \Delta\lambda_i \to 0} \sum_{i=1}^n f(\tau_i) \Delta E(\lambda_i),$$

只要右端极限在依算子范数收敛意义下存在,如通常一样,其中 $\alpha = \lambda_0 < \lambda_1 < \cdots < \lambda_n = \beta$ 是区间 $[\alpha,\beta]$ 的任意分划, $\Delta\lambda_i = \lambda_i - \lambda_{i-1}$, $\tau_i \in [\lambda_{i-1}, \lambda_i]$ 是任取的. 定义

$$\int_{\alpha^-}^\beta f(\lambda) \mathrm{d}E(\lambda) = \lim_{c \uparrow \alpha} \int_c^\beta f(\lambda) \mathrm{d}E(\lambda),$$

只要右端极限存在.

现在终于可以叙述著名的谱分解定理了.

定理 3.4.13 设 $T \in L(H)$ 是自伴算子, m, M 依式(3.4.11), $\{E_\lambda\}$ 是区间 $[m, M]$ 上的谱族,它满足

$$\alpha(E_\beta - E_\alpha) \leqslant T(E_\beta - E_\alpha) \leqslant \beta(E_\beta - E_\alpha), \tag{3.4.17}$$

则成立以下**谱分解公式**

$$T = \int_{m-}^{M} \lambda \mathrm{d}E(\lambda) \qquad (E(\lambda) = E_\lambda); \tag{3.4.18}$$

$$\langle Tx, y\rangle = \int_{m-}^{M} \lambda \mathrm{d}_\lambda \langle E_\lambda x, y\rangle \quad (x, y \in H), \tag{3.4.19}$$

其中 d_λ 表示其后的 $\langle E_\lambda x, y\rangle$ 看作 λ 的函数.

证 显然式(3.4.18)蕴涵式(3.4.19),故只要证式(3.4.18).取定 $a < m$,作分划

$$a = \lambda_0 < \lambda_1 < \cdots < \lambda_n = M$$

令 $\Delta\lambda_i = \lambda_i - \lambda_{i-1}, \delta = \max\limits_{1 \leqslant i \leqslant n} \Delta\lambda_i, \Delta E(\lambda_i) = E(\lambda_i) - E(\lambda_{i-1})$. 其次,任取 $\tau_i \in [\lambda_{i-1}, \lambda_i]$,令 $S = \sum \tau_i \Delta E(\lambda_i)$. 由条件(3.4.17)有

$$\lambda_{i-1}\Delta E(\lambda_i) \leqslant T\Delta E(\lambda_i) \leqslant \lambda_i \Delta E(\lambda_i). \quad (1 \leqslant i \leqslant n) \tag{3.4.20}$$

注意到 $\sum \Delta E(\lambda_i) = I$, 从式(3.4.20)推出

$$\sum_{i=1}^{n} \lambda_{i-1}\Delta E(\lambda_i) \leqslant T \leqslant \sum_{i=1}^{n} \lambda_i \Delta E(\lambda_i).$$

于是

$$S - T = \sum_{i=1}^{n} \tau_i \Delta E(\lambda_i) - \sum_{i=1}^{n} \lambda_{i-1}\Delta E(\lambda_i)$$

$$\leqslant \sum_{i=1}^{n} \Delta\lambda_i \Delta E(\lambda_i) \leqslant \delta \sum_{i=1}^{n} \Delta E(\lambda_i) = \delta I.$$

类似地有 $S - T \geqslant -\delta I$, 于是

$$-\delta I \leqslant S - T \leqslant \delta I,$$

从而 $\| S - T \| \leqslant \delta$ (用定理 3.4.8(vi)). 这正表明式(3.4.18)成立. □

粗略地说,谱分解公式(3.4.18)将自伴算子 T 表成了正投影的(无限)"线性

组合",而正投影是足够简单而容易处理的算子,这就使公式(3.4.18)成为研究自伴算子的方便工具.

然而,你必定已注意到,还遗留一个重大问题:定理 3.4.13 所要求的谱族 $\{E_\lambda\}$ 是否存在? 如何确定它? 这由以下定理解答.

定理 3.4.14 设 $T \in L(H)$ 是自伴算子, $\forall \lambda \in \mathbf{R}, T_\lambda = \lambda I - T, E_\lambda$ 是 H 到 $N(T_\lambda)$ 上的正投影,则 $\{E_\lambda : \lambda \in \mathbf{R}\}$ 是区间 $[m, M]$ 上满足条件(17)的谱族,此处 m, M 依式(3.4.11).

定理 3.4.14 的结论看来是深刻的、并非简单的考虑所能洞察的.但在作了定理3.4.8～命题 3.4.11这一系列准备之后,定理 3.4.14 的证明已不很困难了.尽管如此,我们还是不打算去给出这一证明.你可能更感兴趣的是:一个满足条件(3.4.17)式的谱族其具体形态如何? 要说明这一点,运用具体例子比抽象构造更加有效.

例 3.4.15 (i) 设 $A = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ ①, $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$. 令 $P_i = \mathrm{diag}(0, \cdots, 1, \cdots, 0)$, 其中 1 为对角线上第 i 个元.定义

$$E_\lambda = \sum_{\lambda_i \leqslant \lambda} P_i \quad (\lambda \in \mathbf{R}), \tag{3.4.21}$$

E_λ 必为正投影(用命题 3.4.11(ii)).直接看出: $\lambda \leqslant \mu \Rightarrow E_\lambda \leqslant E_\mu$; $\lambda < \lambda_1 \Rightarrow E_\lambda = 0$ (约定 $\sum_{\varnothing} = 0$), $\lambda \geqslant \lambda_n \Rightarrow E_\lambda = \sum_1^n P_i = I$; 若 $\mu > \lambda$ 且 μ 充分接近于 λ, 则 $E_\mu = E_\lambda$. 可见, $\{E_\lambda\}$ 满足定义 3.4.12 中条件(i)～(iii)(取 $m = \lambda_1, M = \lambda_n$),因而它是区间 $[\lambda_1, \lambda_n]$ 上的谱族.若 $\alpha < \beta$, 则由式(3.4.21)有

$$E_\beta - E_\alpha = \sum_{\alpha < \lambda_i \leqslant \beta} P_i .$$

由此易验证不等式(3.4.17)(换 T 为 A).于是由谱分解公式(3.4.18)有

$$A = \int_{\lambda_1^-}^{\lambda_n} \lambda \mathrm{d}E(\lambda) = \sum_{i=1}^{n} \lambda_i P_i . \tag{3.4.22}$$

当然,在这种简单情况下,你可以直接看出 $A = \sum \lambda_i P_i$, 而根本不必用到谱分解公式(3.4.18).

(ii) 设 T 定义如式(3.4.15),但假定其中 φ 是 $J = [a, b]$ 上的右连续增函数, $\varphi(a) = m, \varphi(b) = M$. 在例 3.4.6(ii)中已指明这样的 T 是 $L^2(J)$ 上的自伴算

① 考虑到有限维 Hilbert 空间中的任何自伴算子在适当法正交基下可表为实对角形,此处的例子对于有限维自伴算子具有一般性.

子,今构成它的一个谱族如下. $\forall\lambda\in\mathbf{R}$, 以 ξ_λ 记集 $\varphi^{-1}[m,\lambda]$ 的特征函数,令

$$E_\lambda u = \xi_\lambda u \quad (u\in L^2(J)).$$

因 ξ_λ 是实值函数,故 E_λ 是自伴算子(参考例 3.4.6(ii));显然 $E_\lambda^2 = E_\lambda$, 因此 E_λ 是正投影. 容易验证: $\lambda\leqslant\mu\Rightarrow E_\lambda\leqslant E_\mu$; $\lambda<m\Rightarrow E_\lambda=0$, $\lambda\geqslant M\Rightarrow E_\lambda = I$; 由 φ 右连续推出,当 $\mu\downarrow\lambda$ 时

$$\|E_\mu u - E_\lambda u\|_2^2 = \int_{\lambda<\varphi(x)\leqslant\mu} |u(x)|^2\mathrm{d}x \to 0 \quad (\forall u\in L^2(J)).$$

因此, $\{E_\lambda\}$ 是 $[m,M]$ 上的一个谱族. 若 $\alpha<\beta$, 则

$$\begin{aligned}\langle T(E_\beta - E_\alpha)u,u\rangle &= \int_{\alpha<\varphi(x)\leqslant\beta}\varphi(x)|u(x)|^2\mathrm{d}x\\ &\leqslant \int_{\alpha<\varphi(x)\leqslant\beta}\beta|u(x)|^2\mathrm{d}x\\ &=\beta\langle(E_\beta - E_\alpha)u,u\rangle \quad (\forall u\in L^2(J)),\end{aligned}$$

这表明 $T(E_\beta - E_\alpha)\leqslant\beta(E_\beta - E_\alpha)$; 同理有 $T(E_\beta - E_\alpha)\geqslant\alpha(E_\beta - E_\alpha)$, 因此条件(3.4.17)式满足. 这样,对于 T 与如上构成的谱族 $\{F_\lambda\}$ 可运用谱分解公式(3.4.18).

那么,如上构成的谱族 $\{E_\lambda\}$ 与由定理 3.4.14 所确定的谱族有何关系呢? $\forall\lambda\in\mathbf{R}$, 如上段末尾所说明的,有 $T_\lambda^- u = (\lambda-\varphi)^- u\,(u\in L^2(J))$, 因而

$$N(T_\lambda^-) = \{u\in L^2(J): 在集\ \varphi^{-1}(\lambda,M)\ 上\ u=0,\text{a.e.}\}.$$

由此不难看出, $E_\lambda u = \xi_\lambda u$ 正是 u 在 $N(T_\lambda^-)$ 上的正投影. 因此,此处所得的 $\{E_\lambda\}$ 与定理 3.4.14 的结论一致.

§3.5 无界算子

无界线性算子较之于有界线性算子,无疑复杂得多,但从应用上看又不能回避. 例如,一些数学物理问题(尤其是量子力学问题)就大量涉及无界线性算子. 本节限于考虑 Hilbert 空间上的无界线性算子,且只介绍某些最基本的概念与结论.

以下设 H 是一复 Hilbert 空间.

§3.5.1 无界算子概念

在无界线性算子理论中,所涉及的算子常常只是在 H 的某个子空间上有定

义.例如,微分算子 $D = \mathrm{d}/\mathrm{d}x$ 显然只定义于 $L^2[0,1]$ 的某个子空间上.对于定义于 H 中的线性算子 T,通常也不必明确指出(或者难以准确描述)其定义域,因此总以 $D(T)$ 表示 T 的定义域.若 $\overline{D(T)} = H$,则说 T 是**稠定的**.通常只有稠定算子才有研究价值,因此设下面给定的算子都是稠定的.要判定一个线性算子 $T: D(T) \subset H \to H$ 是稠定的,并不一定要完全求出定义域 $D(T)$(这并非总是容易的),你只需注意到 $D(T)$ 必为 H 的子空间,若能指出 $D(T)$ 包含 H 的某个基本集,就必定有 $\overline{D(T)} = H$.其次还可注意,若 T 稠定而且有界,则由定理 2.1.7 知 T 可保持范数扩张到 H 上.因此对于有界算子总可以假定它定义在全空间上,而稠定一词可以说是专为无界算子设立的.

如通常所规定的,以 $T \subset S$ 表示算子 S 是 T 的扩张,即 $D(T) \subset D(S)$ 且 $Tx = Sx(\forall x \in D(T))$.

尽管无界算子很不同于有界算子,我们仍然希望按有界线性算子理论的某些思路来展开无界算子理论.下面的定义可与定义 3.4.1 相对照.

定义 3.5.1 设 $T: D(T) \subset H \to H$ 是一线性算子.若对于 $y \in H$,有 $y^* \in H$ 满足恒等式

$$\langle Tx, y\rangle = \langle x, y^*\rangle \quad (\forall x \in D(T)), \tag{3.5.1}$$

则 y^* 由 y 惟一确定(注意 $\overline{D(T)} = H$ 用于此),令 $T^* y = y^*$,称如此确定的映射 $T^*: y \to y^*$ 为 T 的**相伴算子**,以 $D(T^*)$ 记使 $T^* y$ 有定义的 y 之全体.若 $T \subset T^*$,即 T 满足恒等式:

$$\langle Tx, y\rangle = \langle x, Ty\rangle \quad (x, y \in D(T)), \tag{3.5.2}$$

则称 T 为**对称算子**;若 $T = T^*$,则称 T 为**自伴算子**.

若将式(3.5.1)改写成

$$\langle Tx, y\rangle = \langle x, T^* y\rangle \quad (x \in D(T), y \in D(T^*)), \tag{3.5.3}$$

则它正好是式(3.4.2)的推广.与式(3.4.2)不同的是,式(3.5.3)对 x, y 只是**有限制地**使用.定义 3.5.1 中不仅未要求 $D(T^*) = H$,甚至亦未要求 $\overline{D(T^*)} = H$.当然,我们更对那些使 T^* 稠定的算子 T 感兴趣,对称算子显然就是这样的算子.若 $T \in L(H)$,则定义 3.5.1 所定义的 T^* 就是定义 3.4.1 意义下的相伴算子,且当 T 是对称算子时,它就是自伴算子.

直接从(2)推出,若 T 是对称算子,则 $\langle Tx, x\rangle \in \mathbf{R}(\forall x \in D(T))$.

命题 3.4.2 可部分地推广于无界算子.

命题 3.5.2 设 T, S 是 H 上稠定的线性算子.

(i) $\forall \alpha \in \mathbf{C}$,有 $(\alpha T)^* = \bar{\alpha} T^*$;若 $\overline{D(T) \cap D(S)} = H$,则 $T^* + S^* \subset (T + S)^*$,当 $T \in L(H)$ 时包含为等式.

(ii) 若 $\overline{D(TS)} = H$, 则 $S^*T^* \subset (TS)^*$, 当 $T \in L(H)$ 时包含为等式.

(iii) 若 T 是单射且 $\overline{R(T)} = H$, 则 $(T^{-1})^* = (T^*)^{-1}$.

(iv) 若 $\overline{D(T^*)} = H$, 则 $T \subset T^{**}$.

证　(i)(ii)(iv)的证明是直接的. 在(iii)的条件下,有

$$T^*y = 0 \Rightarrow \langle Tx, y\rangle = \langle x, T^*y\rangle = 0 \Rightarrow y = 0,$$

可见 T^* 亦为单射, 因而 $(T^*)^{-1}: R(T^*) \to D(T^*)$ 有定义. 若 $y \in D((T^*)^{-1})$, 则 $\forall x \in D(T)$, 有

$$\begin{aligned}\langle T^{-1}Tx, y\rangle &= \langle x, y\rangle = \langle x, T^*(T^*)^{-1}y\rangle \\ &= \langle Tx, (T^*)^{-1}y\rangle,\end{aligned}$$

这推出 $(T^{-1})^*y = (T^*)^{-1}y$. 其次设 $y \in D((T^{-1})^*)$, 则 $\forall x \in D(T)$, 有

$$\langle x, y\rangle = \langle T^{-1}Tx, y\rangle = \langle Tx, (T^{-1})^*y\rangle,$$

这与(1)对照得出 $y = T^*(T^{-1})^*y \in R(T^*) = D((T^*)^{-1})$, 这就证得 $(T^{-1})^* = (T^*)^{-1}$. □

对于无界算子无连续性可用,这无疑是一重大缺陷. 不过,这可由考虑“闭性”而得到一定弥补. 如在 §2.4 中已提到的, 约定称 $G(T) \triangleq \{(x, Tx): x \in D(T)\}$ 为 T 的图像, 当 $G(T)$ 是 $H \times H$ 的闭子集时称 T 为**闭线性算子**, 简称**闭算子**. 显然, T 是闭算子相当于(参看式(2.4.2))

$$\left.\begin{array}{l} x_n \to x \\ Tx_n \to y \end{array}\right\} \Rightarrow x \in D(T) \text{ 且 } Tx = y. \tag{3.5.4}$$

若 $T \subset T_1$, T_1 是闭算子, 则称 T_1 为 T 的**闭扩张**. T 的最小闭扩张称为 T 的**闭包**, 记作 $\bar{T}$. 显然, T 是闭算子 $\Leftrightarrow T = \bar{T}$. 闭线性算子之所以特别重要, 其理由在于:应用上重要的线性算子几乎都是闭线性算子,或者有自然的闭扩张. 有界线性算子则并不具有这种普遍性.

关于闭算子有以下结论.

命题 3.5.3　设 T 是 H 上稠定的线性算子.

(i) 若 T 是闭算子, 则 T 有界 $\Leftrightarrow D(T) = H$.

(ii) T^* 必为闭算子;特别, 自伴算子是闭算子.

(iii) 若 T 是对称算子, 则闭包 $\bar{T}$ 存在且亦为对称算子.

证　(i) 设 T 有界, $\forall x \in H$, 取 $\{x_n\} \subset D(T)$, 使 $x_n \to x (n \to \infty)$. $\{Tx_n\}$ 必为 Cauchy 列, 故可设 $Tx_n \to y (n \to \infty)$, 于是由条件(3.5.4)推出 $x \in D(T)$. 因此 $D(T) = H$. 反之, 当 $D(T) = H$ 时由闭图像定理推出 T 有界.

(ii) 设 $y_n \to y, T^* y_n \to z(n \to \infty)$, 则从

$$\langle Tx, y_n \rangle = \langle x, T^* y_n \rangle \quad (\forall x \in D(T))$$

推出 $\langle Tx, y \rangle = \langle x, z \rangle (\forall x \in D(T))$, 这表明 $y \in D(T^*)$ 且 $z = T^* y$. 因此 T^* 是闭算子.

(iii) 若 T_1 是 T 的闭扩张,则必定 $\overline{G(T)} \subset G(T_1)$. 因此,若 $\overline{G(T)}$ 是某个线性算子 $\bar{T}$ 的图像,则 $\bar{T}$ 必为 T 的闭包. 为证有线性算子 $\bar{T}$ 使 $\overline{G(T)} = G(\bar{T})$, 只需说明当 $(x, y_i) \in \overline{G(T)} (i = 1,2)$ 时, $y_1 = y_2$. 取 $\{x_n^i\} \subset D(T)$, 使 $x_n^i \to x$, $Tx_n^i \to y_i (n \to \infty, i = 1,2)$, 则 $\forall z \in D(T)$, 有

$$\begin{aligned}\langle y_1 - y_2, z \rangle &= \lim_n \langle Tx_n^1 - Tx_n^2, z \rangle \\ &= \lim_n \langle x_n^1 - x_n^2, Tz \rangle = 0, \quad (\text{用对称性})\end{aligned}$$

这得出 $y_1 = y_2$. 因此 $\overline{G(T)} = G(\bar{T})$, $\bar{T}$ 是 T 的闭包. $\forall x, y \in D(\bar{T})$, 必有 x_n, $y_n \in D(T)(n \in \mathbf{N})$, 使得 $x_n \to x, Tx_n \to \bar{T}x, y_n \to y, Ty_n \to \bar{T}y(n \to \infty)$. 于是,由

$$\langle Tx_n, y_n \rangle = \langle x_n, Ty_n \rangle \quad (n \in \mathbf{N})$$

推出 $\langle \bar{T}x, y \rangle = \langle x, \bar{T}y \rangle$. 这正表明 $\bar{T}$ 是对称算子. □

若不假定 T 稠定,则命题 3.5.3 中的结论(i)应改为:若 T 是闭算子,则 T 有界 $\Leftrightarrow D(T)$ 是闭子空间.

§3.5.2 谱理论

关于无界算子的中心课题亦是谱理论. 以下设 T 是 H 上稠定的线性算子,首先推广谱概念如下(参照定义 3.1.2).

定义 3.5.4 设 $T_\lambda = \lambda I - T$, 令

$$\rho(T) = \{\lambda \in \mathbf{C} : T_\lambda^{-1} \text{ 有定义且 } T_\lambda^{-1} \in L(H)\}; \tag{3.5.5}$$

$$\sigma_p(T) = \{\lambda \in \mathbf{C} : N(T_\lambda) \neq \{0\}\}; \tag{3.5.6}$$

$$\sigma_c(T) = \{\lambda \in \mathbf{C} : T_\lambda^{-1} \text{ 稠定但 } D(T_\lambda^{-1}) \neq H\}; \tag{3.5.7}$$

$$\sigma_r(T) = \{\lambda \in \mathbf{C} : T_\lambda^{-1} \text{ 有定义但非稠定}\}. \tag{3.5.8}$$

以上 4 个集分别称为 T 的**正则值集**、**点谱**、**连续谱**与**剩余谱**,而称

$$\sigma(T) = \sigma_p(T) \cup \sigma_c(T) \cup \sigma_r(T) = \mathbf{C} \setminus \rho(T)$$

为 T 的**谱**. 称 $R(\lambda, T) = T_\lambda^{-1} (\lambda \in \rho(T))$ 为 T 的**预解式**.

下面的结果推广了有界自伴算子的一些熟知性质.

定理 3.5.5　设 T 是自伴算子,则以下结论成立:

(i) $\forall \lambda \in \mathbf{C}, \lambda \in \rho(T) \Leftrightarrow \exists \beta > 0$, 使得

$$\| T_\lambda x \| \geqslant \beta \| x \| \quad (\forall x \in D(T)). \tag{9}$$

(ii) $\sigma(T) \subset \mathbf{R}$ 且 $\sigma(T)$ 是闭集(参照定理 3.4.5(ii)与定理 3.1.3).

证　(i) 若 $\lambda \in \rho(T)$, 则由命题 2.1.8 推出有 $\beta > 0$ 使式(3.5.9)成立. 反之,若有 $\beta > 0$ 使(3.5.9)成立,则由命题 2.1.8 知 T_λ 有有界逆. 因 T (从而 T_λ)是闭算子(命题 3.5.3(ii)),故 T_λ^{-1} 亦为闭算子,从而 $R(T_\lambda)$ 是闭集(参看命题 3.5.3 证明之后的说明). 为证 $\lambda \in \rho(T)$, 只要指明 $R(T_\lambda) = H$; 为此又只要说明 $R(T_\lambda)^\perp = \{0\}$ (对照定理 3.4.5(ii)之证). 设 $y \in R(T_\lambda)^\perp$, 则

$$0 = \langle T_\lambda x, y \rangle = \langle \lambda x, y \rangle - \langle Tx, y \rangle$$

$$= \langle x, \bar{\lambda} y - Ty \rangle \quad (\forall x \in D(T)),$$

这推出 $\bar{\lambda} y = Ty$, 因而

$$\bar{\lambda} \| y \|^2 = \langle Ty, y \rangle = \langle y, Ty \rangle = \lambda \| y \|^2.$$

这推出 $y = 0$ (否则 $\lambda = \bar{\lambda}, 0 \neq y \in N(T_\lambda)$, 与条件(3.5.9)矛盾),如所要证.

(iii) 首先证 $\sigma(T) \subset \mathbf{R}$. 设 $\lambda = \alpha + \mathrm{i}\beta \in \sigma(T), \alpha, \beta \in \mathbf{R}$, 则 $\forall x \in D(T)$, 有

$$2\mathrm{i}\,\mathrm{Im}\langle T_\lambda x, x \rangle = \langle T_\lambda x, x \rangle - \langle x, T_\lambda x \rangle = 2\mathrm{i}\beta \| x \|^2;$$

$$\beta \| x \|^2 = \mathrm{Im}\langle T_\lambda x, x \rangle \leqslant \| T_\lambda x \| \| x \|.$$

这推出 $\| T_\lambda x \| \geqslant \beta \| x \|$. 因 $\lambda \bar{\in} \rho(T)$, 故用已证之(i)得 $\beta \leqslant 0$. 以 $-T$ 换 T 得 $\beta \geqslant 0$, 因此 $\beta = 0, \lambda \in \mathbf{R}$.

其次证 $\rho(T)$ 为开集(从而 $\sigma(T)$ 为闭集). 任给 $\lambda_0 \in \rho(T), \lambda \in \mathbf{C}$, 有

$$\| T_\lambda x \| \geqslant \| T_{\lambda_0} x \| - | \lambda - \lambda_0 | \| x \|$$

$$\geqslant (\| T_{\lambda_0}^{-1} \|^{-1} - | \lambda - \lambda_0 |) \| x \| \quad (x \in D(T)),$$

可见当 λ 充分邻近 λ_0 时有 $\lambda \in \rho(T)$ (用(i)),故 $\rho(T)$ 为开集. □

注意,对于无界自伴算子 T, 并不能断定 $\sigma(T)$ 是紧集,你很快就会看到这一点(参看命题 3.5.9 与 3.5.10).

在上节中,我们曾将自伴算子比拟于实数,而将 U 算子比拟于绝对值为 1 的复数. 现在利用这一比拟引导出一个重要变换. 在复变函数论中证明了,分式线性变换

$$z = (x - \mathrm{i})/(x + \mathrm{i}) \quad (x \in \mathbf{R})$$

将实直线变换为单位圆周.类比于此,自然会问:可否断定变换

$$U = (T - \mathrm{i}I)(T + \mathrm{i}I)^{-1} \tag{3.5.10}$$

将自伴算子 T 变为 U 算子 U? 如(10)所定义的 U 称为 T 的 **Cayley 变换**.下面的定理为上述猜测给出肯定回答.

定理 3.5.6 设 T 是 H 上的自伴算子, U 如(10),则 U 是 U 算子, $1 \notin \sigma_p(U)$, 且

$$T = \mathrm{i}(I + U)(I - U)^{-1}. \tag{3.5.11}$$

证 因 $\pm \mathrm{i} \in \rho(T)$ (用定理 3.5.5(ii)),故 $(T \pm \mathrm{i}I)^{-1} \in L(H)$, 因而式(3.5.10)表出一个线性算子 $U: H \to H, R(U) = R(T - \mathrm{i}I) = H$. $\forall x \in H$, 令 $y = (T + \mathrm{i}I)^{-1}x$, 则

$$\begin{aligned} \| Ux \|^2 &= \langle Ty - \mathrm{i}y, Ty - \mathrm{i}y \rangle \\ &= \langle Ty + \mathrm{i}y, Ty + \mathrm{i}y \rangle \qquad (\text{用式}(3.5.2)) \\ &= \| (T + \mathrm{i}I)y \|^2 = \| x \|^2, \end{aligned}$$

可见 U 是 U 算子(命题 3.4.4(iii)).设 x, y 依然如上式,则

$$x = (T + \mathrm{i}I)y, \quad Ux = (T - \mathrm{i}I)y;$$

由此得

$$(I + U)x = 2Ty, \quad (I - U)x = 2\mathrm{i}y. \tag{3.5.12}$$

因 $(I - U)x = 0 \Rightarrow y = 0 \Rightarrow x = 0$, 故 $1 \notin \sigma_p(U)$, 从而 $(I - U)^{-1}$ 定义于

$$R(I - U) = R((T + \mathrm{i}I)^{-1}) = D(T)$$

上.由式(3.5.12)有

$$\begin{aligned} Ty &= \frac{1}{2}(I + U)(I - U)^{-1}(2\mathrm{i}y) \\ &= \mathrm{i}(I + U)(I - U)^{-1}y, \end{aligned}$$

这表明式(3.5.11)成立. □

至此,我们尚未触及谱理论的中心问题:谱分解.不过,已建立的结果为谱分解问题的解决准备了条件.尽管为达到最终结论仍然有一段艰难的路程,以至在这个导引性介绍中,我们已不宜深入下去了,但基本的思路还是很简单的,概括说来就是:

(i) 每个 U 算子 U 必可表成 $U = \mathrm{e}^{iA}$, A 是某个有界自伴算子.

(ii) 利用公式 $U = \mathrm{e}^{iA}$, 可从有界自伴算子的谱分解(参看定理 3.4.13)得出 U 算子的谱分解.

(iii) 利用公式(11),可从 U 算子的谱分解得出无界自谱算子的谱分解.

当然,以上三件事实际做起来都不太简单,但最终可供应用的结论是简单明晰的,它们就是

定理 3.5.7　设 $U \in L(H)$ 是 U 算子,则存在区间 $[-\pi,\pi]$ 上的谱族 $\{E_\theta\}$, 使得

$$U = \int_{-\pi}^{\pi} \mathrm{e}^{i\theta}\mathrm{d}E(\theta), \quad E(\theta) = E_\theta,$$

其中积分是依算子范数收敛的 RS 积分.

定理 3.5.8　设 T 是 H 上稠定的自伴算子, U 是 T 的 Cayley 变换, $\{E_\theta\}$ 是 U 算子 $-U$ 的谱族(依定理 3.5.7), $F_\lambda = E_{2\arctan\lambda}$, 则

$$\langle Tx, x\rangle = \int_{-\infty}^{\infty} \lambda \mathrm{d}_\lambda \langle F_\lambda x, x\rangle \quad (x \in D(T)).$$

§3.5.3　乘法算子与微分算子

现在考虑无界自伴算子的两个典型例子,它们在量子力学中是不可缺少的工具.

取 $H = L^2(\mathbf{R})$. 设 φ 是 $\mathbf{R}$ 上的实可测函数,它在 $\mathbf{R}$ 上局部有界,这意味着在任何有限区间上有界,这就使得对任给 $u \in L^2(\mathbf{R})$, φu 在任何有限区间上平方可积. 利用 φ 定义如下**乘法算子**(参照式(3.4.15)):

$$Tu = \varphi u \quad (u \in D(T)), \tag{3.5.13}$$

其定义域

$$D(T) = \{u \in L^2(\mathbf{R}) : \varphi u \in L^2(\mathbf{R})\}.$$

对任何有限区间 $\delta \subset \mathbf{R}$, 显然有 $\chi_\delta \in D(T)$. 因这样的 χ_δ 构成 $L^2(\mathbf{R})$ 的一个基本集(参看例 1.3.8(iv)),故 T 在 $L^2(\mathbf{R})$ 上稠定. 除非 $\varphi \in L^\infty(\mathbf{R})$, T 一般是无界的. 例如,取 $\varphi(x) \equiv x$, $u_n = \chi_{[0,n]}$, 则

$$\| u_n \|_2^2 = \int_0^n \mathrm{d}x = n,$$

$$\| Tu_n \|_2^2 = \int_0^n x^2 \mathrm{d}x = \frac{n^3}{3},$$

$$\frac{\|Tu_n\|_2}{\|u_n\|_2}=\frac{n}{\sqrt{3}}\to\infty\quad(n\to\infty).$$

关于乘法算子(3.5.13)的基本结论是.

命题 3.5.9 T 是 $L^2(\mathbf{R})$ 上的无界自伴算子;若 $\varphi(x)\equiv x$, 则 $\sigma(T)=\mathbf{R}$, $\sigma_p(T)=\varnothing$.

证 任给 $u,v\in D(T)$, 显然有

$$\langle Tu,v\rangle=\langle\varphi u,v\rangle=\langle u,\varphi v\rangle=\langle u,Tv\rangle,$$

可见 T 是对称算子.另一方面,若 $v\in D(T^*)$, 则由式(3.5.3)有

$$\langle u,T^*v\rangle=\langle Tu,v\rangle=\int_{\mathbf{R}}u(x)\overline{\varphi(x)v(x)}\mathrm{d}x\quad(u\in D(T)).\tag{3.5.14}$$

任给 $n\in\mathbf{N}$, 定义 $L^2[-n,n]$ 上的有界线性泛函

$$f_n(u)=\int_{-n}^{n}u(x)\overline{\varphi(x)v(x)}\mathrm{d}x.$$

因每个 $u\in L^2[-n,n]$ 可看作 $D(T)$ 中的元(在 $[-n,n]$ 之外令 $u(x)=0$),故依式(3.5.14)有

$$|f_n(u)|=|\langle u,T^*v\rangle|\leqslant\|u\|_2\|T^*v\|_2.$$

于是

$$\left[\int_{-n}^{n}|\varphi(x)v(x)|^2\mathrm{d}x\right]^{1/2}=\|f_n\|\leqslant\|T^*v\|_2,\quad(\text{用定理 }2.3.6)$$

这推出 $\|\varphi v\|_2\leqslant\|T^*v\|_2$, 即 $\varphi v\in L^2(\mathbf{R})$, $v\in D(T)$. 故 T 是自伴算子.因此由定理 3.5.5(ii)有 $\sigma(T)\subset\mathbf{R}$.

以下设 $\varphi(x)\equiv x$. 今证 $\sigma(T)=\mathbf{R}$. 若此结论不真,则有实数 $\lambda\overline{\in}\sigma(T)$. 于是对任给 $v\in L^2(\mathbf{R})$, 方程

$$\lambda u-Tu=v$$

有惟一解 $u\in D(T)$, 这就推出

$$u(x)\triangleq(\lambda-x)^{-1}v(x)\in L^2(\mathbf{R}).$$

但以上结论显然对 $v=\chi_{[\lambda,\lambda+1]}\in L^2(\mathbf{R})$ 就不能成立,故有 $\sigma(T)=\mathbf{R}$.

最后证 T 无特征值.任给 $\lambda\in\mathbf{R}$. 若有 $u\in D(T)$, 使得 $Tu=\lambda u$, 则由

$$(x-\lambda)u(x)=0.\text{a.e.}$$

推出 $u(x)=0$, a.e.,可见 $\lambda \overline{\in} \sigma_p(T)$. 因此 $\sigma_p(T)=\varnothing$. □

从形式上看,乘法算子 $Tu(x)=xu(x)$ 的定义简单得近于平凡.你可能会感到困惑:这样一个毫不起眼的算子如何会有益于量子力学!简单说来事情是这样的:给定 $u\in D(T)$,当 $\|u\|_2=1$ 时,$|u|^2$ 可看作一概率密度函数,它描述了一个在实轴上运动的质点(如某种粒子)位置的统计规律.任给可测集 $A\subset\mathbf{R}$,

$$\int_A |u(x)|^2\mathrm{d}x$$

就是质点位于集 A 中的概率;而期望值

$$\int_{\mathbf{R}} x|u(x)|^2\mathrm{d}x=\langle xu,u\rangle=\langle Tu,u\rangle$$

则表示质点的期望位置或平均坐标.这就表明,Tu 在统计平均的意义上刻画了质点的位置,而位置的方差是

$$\int_{\mathbf{R}}(x-\mu)^2|u(x)|^2\mathrm{d}x=\langle (T-\mu I)^2u,u\rangle,$$

其中 $\mu=\langle Tu,u\rangle$. 鉴于以上解释,T 也称为**位置算子**或**坐标算子**.如所周知,描述处于高速运动中的微观粒子的准确位置,是毫无意义的;其状态只能用在一定区域(或区间)内出现的概率来刻画.正是这一实际上很简单的思想导致了产生量子力学的物理学革命,而所依据的数学工具就是如坐标算子一类的无界线性算子.

对于乘法算子 T,Cayley 变换有很简单的表示.设 T 与 U 分别依式(3.5.13)与(3.5.10),则

$$Uu=zu,\quad z=\frac{\varphi-\mathrm{i}}{\varphi+\mathrm{i}}\quad(u\in L^2(\mathbf{R})).$$

因 $|z|\equiv 1$,易验证 U 为 U 算子(参考例 3.4.6(ii)).从等式

$$(\varphi+\mathrm{i})zu=(\varphi-\mathrm{i})u$$

解出

$$\begin{aligned}Tu=\varphi u&=\frac{\mathrm{i}(1+z)}{1-z}u\\&=\mathrm{i}(I+U)(I-U)^{-1}u\quad(u\in L^2(\mathbf{R})),\end{aligned}$$

这正好与公式(3.5.11)相合.

另一个同样重要的**算子**是**微分算子**:

$$Du = \mathrm{i}u' \quad (u \in \mathscr{D}), \tag{3.5.15}$$

其定义域

$$\mathscr{D} = \{u \in L^2(\mathbf{R}) : u \text{ 在有限区间上绝对连续且 } u' \in L^2(\mathbf{R})\}.$$

显然 $C_c^1(\mathbf{R}) \subset \mathscr{D}$，而 $C_c^1(\mathbf{R})$ 在 $L^2(\mathbf{R})$ 中稠密(参看例 1.3.8(v))，故微分算子 D 在 $L^2(\mathbf{R})$ 中稠定. 如同位置算子 T 一样，D 也是无界的. 取

$$u_n(x) = (1 - n|x|)\chi_n(x),$$

χ_n 是区间 $(-1/n, 1/n)$ 的特征函数，则 $u_n \in \mathscr{D}$，

$$\|u_n\|_2^2 = 2\int_0^{1/n} (1 - nx)^2 \mathrm{d}x = \frac{2}{n}\int_0^1 y^2 \mathrm{d}y = \frac{2}{3n},$$

$$\|Du_n\|_2^2 = 2\int_0^{1/n} n^2 \mathrm{d}x = 2n,$$

$$\|Du_n\|_2 / \|u_n\|_2 = \sqrt{3}n \to \infty \quad (n \to \infty).$$

尽管微分算子 D 的定义式如同位置算子 T 一样简单，但它涉及微分运算，因而对它的分析要复杂一些. 不过，仍可建立如同命题 3.5.9 一样的结论：

命题 3.5.10 微分算子(3.5.15)是 $L^2(\mathbf{R})$ 上的无界自伴算子；$\sigma(D) = \mathbf{R}$；$\sigma_p(D) = \varnothing$.

在量子力学中，$-(h/2\pi)D$ 是所谓**动量算子**，此处 h 是 Plank 常数. 这就决定了微分算子 D 在量子力学中的重要位置. 关于它的进一步的讨论可参考文献[1]第 11 章.

有趣的是，初看起来似不相关的位置算子 T 与微分算子 D 有密切的联系. 为简单起见，下面仅给出一个启发性的推导，严格的讨论需要更多的分析知识. 首先指出：$C_c^1(\mathbf{R}) \subset D(T) \cap \mathscr{D}$，且 $C_c^1(\mathbf{R})$ 在 $L^2(\mathbf{R})$ 中稠密. 任给 $u \in C_c^1(\mathbf{R})$，以 $\hat{u}$ 记 u 的 Fourier 变换(参看例 1.3.9)，并令 $Fu = \hat{u}$，则

$$DFu(x) = \mathrm{i}\hat{u}'(x) = \mathrm{i}\int_{\mathbf{R}} \mathrm{e}^{-\mathrm{i}xy} u(y)\mathrm{d}y$$

$$= \int_{\mathbf{R}} \mathrm{e}^{-\mathrm{i}xy} y u(y)\mathrm{d}y$$

$$= (yu)^{\wedge}(x) = FTu(x);$$

$$FDu(x) = \int_{\mathbf{R}} \mathrm{e}^{-\mathrm{i}xy} \mathrm{i}u'(y)\mathrm{d}y$$

$$=\int_{\mathbf{R}} x\mathrm{e}^{-\mathrm{i}xy}u(y)\mathrm{d}y \qquad \text{(分部积分)}$$

$$= x\hat{u}(x) = TFu(x).$$

这就表明,限制在 $C_c^1(\mathbf{R})$ 上考虑,成立等式

$$DF = FT, \quad FD = TF;$$

或

$$D = FTF^{-1} = F^{-1}TF.$$

如果运用 $L^2(\mathbf{R})$ 上的 Fourier 变换,可得出以上结论的适当推广,但我们已无法在此深入讨论了.

习　题

1. 设 $T,S \in L(X)$, $\|S\| < \|T^{-1}\|^{-1}$, 证明

$$\|(T+S)^{-1} - T^{-1}\| \leqslant \frac{\|T^{-1}\|^2\|S\|}{1-\|T^{-1}\|\|S\|}.$$

2. 设 $|\lambda| > \|T\|$, 证明 $\|R(\lambda,T) - \sum_0^n \lambda^{-k-1}T^k\| \leqslant \|\lambda^{-1}T\|^{n+1}/(|\lambda| - \|T\|)$.

3. 设 $T,S \in L(X)$, 证明 $r_\sigma(TS) = r_\sigma(ST)$.

4. 证明: $\sigma(T) = \{0\} \Leftrightarrow \|\lambda^n T^n\| \to 0(n \to \infty, \forall \lambda \in \mathbf{C})$.

5. 设 $Tx = (0,x_1,x_2,\cdots)$, 证明 $T \in L(l^p)(1 \leqslant p < \infty)$, $\sigma_p(T) = \varnothing$.

6. 设 $Tx = (x_2,x_3,\cdots)$, $T \in L(l^p)(1 \leqslant p < \infty)$, 求 $\sigma_p(T)$, $\sigma(T)$.

7. 任给非空紧集 $\sigma \subset \mathbf{C}$, 作 $T \in L(l^1)$, 使得 $\sigma(T) = \sigma$.

8. 设 $T \in L(C[0,1])$ 定义为 $Tu(x) = xu(x)$, 求 $\sigma(T)$ 与 $\sigma_p(T)$.

9. 设 $T \in L(C(J))$ 定义为 $Tu = \varphi u$, $\varphi \in C(J)$ 是给定的,求 $\sigma(T)$.

10. 设 $\lambda \in \mathbf{C}$, $v \in C(J)$, $J = [0,1]$, 证明方程 $u(x) = \lambda\int_0^x u(y)\mathrm{d}y + v(x)$ 有惟一解 $u \in C(J)$.

11. 固定 $\lambda \in \rho(T)$, 设 $f(\tau) = (\lambda - \tau)^{-1}$, 证明 $f(T) = R(\lambda,T)$.

12. 设 $\lambda \in \rho(T) \cap \rho(S)$, 证明 $R(\lambda,T) - R(\lambda,S) = R(\lambda,T)(T - S)R(\lambda,S)$.

13. 证明 $\sigma(R(\lambda,T)) = \{(\lambda - \tau)^{-1} : \tau \in \sigma(T)\}$.

14. 设 $\mu \in \rho(R(\lambda,T))$, 证明 $R(\mu,R(\lambda,T)) = \mu^{-1}I + \mu^{-2}R(\lambda - \mu^{-1},T)$.

15. 设 $f(T) = 0$, 证明 $\sigma(T) \subset \{\lambda \in \mathbf{C} : f(\lambda) = 0\}$.

16. 证明: $\forall y \in X$, $f(T)x = y$ 有惟一解 $\Leftrightarrow \forall \lambda \in \sigma(T)$, 有 $f(\lambda) \neq 0$.

17. 设 $\forall \lambda \in \sigma(T)$, 有 $\lambda \overline{\in} (-\infty,0]$, 证明存在 $A \in L(X)$, 使得 $A^2 = T$.

18. 设 $\forall \lambda \in \sigma(T)$, 有 $\lambda \overline{\in} (-\infty,0]$, 证明存在 $A \in L(X)$, 使得 $\mathrm{e}^A = T$.

19. 设 $T \in \mathrm{CL}(l^2)$ 定义为 $Tx = (x_{i+1}/i)$，求 $\sigma(T)$ 与 $\sigma_p(T)$.

20. 设 $T \in \mathrm{CL}(l^2)$ 定义为 $Tx = (x_i/i)$，求 $\sigma(T)$ 与 $\sigma_p(T)$.

21. 设 $Tx = (\alpha_i x_i)$，$\{\alpha_i\} = [0,1]$，证明 $T \in L(l^2) \setminus \mathrm{CL}(l^2)$.

22. 研究积分方程 $u(x) - \lambda\int_0^1 e^{x-y}u(y)\mathrm{d}y = v(x)$，$v \in C[0,1]$ 是已知的.

23. 设 $f(\lambda)$ 在 $\mathbf{C}$ 上解析，$f(0) = 0$，$T \in \mathrm{CL}(X)$，证明 $f(T) \in \mathrm{CL}(X)$.

24. 设 $T \in L(H)$，证明 $N(I + T^* T) = \{0\}$.

25. 设 $T \in L(H)$，$\| T \| \leqslant 1$，证明 $N(I - T) = N(I - T^*)$.

26. 设 $T \in L(H)$，$|\langle Tx, x\rangle| \geqslant m\|x\|^2 (\forall x \in H)$，$m > 0$，证明 T 可逆.

27. 设 $T \in L(H)$，$|\alpha| = |\beta| = 1$，证明 $\alpha T + \beta T^*$ 是正规算子.

28. 设 T 是正规算子，证明 T 可逆 $\Leftrightarrow \exists \beta > 0$，使 $\| Tx \| \geqslant \beta\| x \| (\forall x \in H)$.

29. 设 $T = T^* \in L(H)$，证明 $N(T) = R(T)^{\perp}$.

30. 设 $T = T^* \in L(H)$，$S \in L(H)$，$TS = I$，证明 T 可逆.

31. 设 $T = T^* \in L(H)$，$\mathrm{Im}\lambda \neq 0$，证明 $\| T_\lambda^{-1} \| \leqslant 1/|\mathrm{Im}\lambda|$.

32. 设 $T = T^* \in \mathrm{CL}(H)$，$H \neq \{0\}$，证明 $\sigma_p(T) \neq \varnothing$.

33. 证明：$T^* = -T \Leftrightarrow \mathrm{Re}\langle Tx, x\rangle = 0 (\forall x \in H)$.

34. 设 $T^* = -T \in L(H)$，$\sigma(T)$ 有何特征?

35. 证明：$P \in L(H)$ 是正投影 $\Leftrightarrow \langle Px, x\rangle = \| Px \|^2$.

36. 设 $P \in L(H)$ 是正投影，证明 $Px = x \Leftrightarrow \| Px \| = \| x \|$.

37. 设 $T: H \to H$ 是线性的，$\langle Tx, y\rangle = \langle x, Ty\rangle$，证明 $T \in L(H)$.

38. 设 $T: D(T) \subset L^2(\mathbf{R}) \to L^2(\mathbf{R})$ 是乘法算子，即 $Tu = \varphi u$，此处 $\varphi \in C^1(\mathbf{R})$ 是给定的，假定 φ 有闭值域. 证明 $\sigma(T) = \varphi(\mathbf{R})$.

第四章　非线性算子与非线性泛函

至此为止,你所接触到的泛函分析知识几乎都是关于线性算子的.实际上,非线性问题是一个更为广阔且更具挑战性的领域,与其巨大的多样性与复杂性相比,线性算子理论不过是一个序章而已.如果说,线性算子理论在一定程度上可看作"无限维空间上的线性代数学",那么,非线性算子理论则更像"无限维空间上的分析学".对于后者,自然首先要考虑到借鉴经典的数学分析,有时也要借鉴线性算子理论.这两方面的借鉴或者仿效常常是很有效的,在非线性算子理论发展的初始阶段尤其如此.当然,深入的展开越来越依赖于新观念的创设与新方法的开发.然而,这些多半已越出本书的范围之外.有兴趣的读者可阅读非线性泛函分析方面的专著.

对于非线性算子的研究,虽然已积累了大量的材料,但与已相当系统且完整的线性算子理论不同,它们更像一些独立的篇章,难以融汇于某一公认的体系.如果你看到本章这个导引性的介绍似乎是一些互不相属的内容的拼合,将不应感到奇怪.

在本章中,X,Y,Z 等记实 Banach 空间、F,G 等记非线性算子,f,g 等记非线性泛函,Ω 通常记某个非空开集.

§4.1　微分理论

本节给出 Banach 空间上微分学的基本框架.你会看到,它几乎是经典微分学在更高层次上的重建.

§4.1.1　Fréchet 导数

如同在有限维空间中一样,微分概念源于非线性函数的**局部线性化**.给定开集 $\Omega\subset X$.对于一个算子 $F:\Omega\to Y$ 与 $x_0\in\Omega$,我们常常希望在 x_0 邻近有近似公式

$$Fx\approx Fx_0+T(x-x_0),$$

其中 T 是某个线性算子.这引导出以下定义.

定义 4.1.1　设 $F:\Omega\subset X\to Y$,$x_0\in\Omega$.若存在 $T\in L(X,Y)$,使得当 $h\in X$,$\|h\|\to 0$ 时有

$$\| F(x_0+h) - Fx_0 - Th \| = o(\| h \|), \tag{4.1.1}$$

即

$$\lim_{\|h\|\to 0} \| F(x_0+h) - Fx_0 - Th \| / \| h \| = 0,$$

则说 F 在 x_0 处 **Fréchet 可微或 F 可微**，而称 T 为 F 在 x_0 的 **Fréchet 导数或 F 导数**，记作 $F'(x_0)$.

当不同时涉及其他类型的微分时，分别将"Fréchet 可微"与"Fréchet 导数"简称为"可微"与"导数"；导数也称为导算子.

直接看出，条件(1)相当于：$\forall \varepsilon > 0, \exists \delta > 0$，使得

$$\| h \| \leqslant \delta \Rightarrow \| F(x_0+h) - Fx_0 - F'(x_0)h \| \leqslant \varepsilon \| h \|. \tag{4.1.2}$$

式(4.1.1)的更简略的写法是

$$F(x_0+h) - Fx_0 = F'(x_0)h + o(h). \tag{4.1.1$'$}$$

由式(4.1.1)直接推出：当 $F'(x_0)$ 存在时，它必然是惟一确定的，且 F 在 x_0 连续. 若 F 在 Ω 内处处可微，则得到一个"导函数"：

$$F' : \Omega \subset X \to L(X,Y), \quad x \to F'(x).$$

当 $F'(x)$ 的导数存在时记作 $F''(x)$，称它为 F 的 2 阶导数. 依此类推，可定义 F 的 n 阶导数 $F^{(n)}(x)$. 当 $F^{(n)}(x)$ 在 Ω 内处处存在且对 x 连续时，称 F 为 **C^n 映射**或 **C^n 函数**，记作 $F \in C^n$. 约定 $F \in C^0$ 意味着 F 连续，$F^{(0)} = F$；$F \in C^\infty \Leftrightarrow F \in C^n (\forall n \geqslant 0)$. 上述所有概念与记号与经典微分学并无差别. 惟一不同的是，现在"导数" $F'(x)$ 并非是一个数，而是一个算子；而且 $F, F', F'', \cdots$ 取值于互不相同的空间中，通常可用的式子 $F(x) - F'(x)$ 等一般不再有意义. 这些新的情况将是某些技术性困难的根源.

虽然定义 4.1.1 形式上很简单，但对于无限维空间上的函数的导数，我们还没有任何直观印象，这应当从具体例子中获得.

例 4.1.2 (i) 设 $x(\cdot) : J \subset \mathbf{R} \to X$，$J$ 是一区间. 容易验证，

$$L(\mathbf{R}, X) \to X, \quad T \to T(1)$$

是一等距同构，因而不妨将 $T \in L(\mathbf{R}, X)$ 等同于 $x_0 = T(1) \in X$ (注意 $T(t) = tx_0, t \in \mathbf{R}$). 显然

$$x(t+h) - x(t) = hx_0 + o(h) \quad (h \to 0)$$

的充要条件是 $h^{-1}[x(t+h) - x(t)] \to x_0 (h \to 0)$，这正表明

$$x'(t) = x_0 = \lim_{h\to 0}\frac{x(t+h)-x(t)}{h}.$$

可见,当 $x'(t)$ 存在时,它就是式(2.4.7)所定义的导数.

(ii) 设 H 是实 Hilbert 空间, $f(x) = \|x\|, 0 \neq x \in H$, 则当 $h \to 0$ 时,

$$\begin{aligned} f(x+h) - f(x) &= \frac{\|x+h\|^2 - \|x\|^2}{\|x+h\| + \|x\|} \\ &= \frac{2\langle x,h\rangle + \|h\|^2}{\|x+h\| + \|x\|} \\ &= \langle \|x\|^{-1}x, h\rangle + o(\|h\|). \end{aligned}$$

注意到 $H^* = H$(定理2.3.6),将上式与式(4.1.1)′对照得出 $f'(x) = x/\|x\|$.

一般地,若 $f:\Omega\subset X\to \mathbf{R}$ 可微, $x\in\Omega$, 则 $f'(x)\in X^*$, 称它为 f 在 x 的**梯度**,通常写作 $\nabla f(x)$. 因此,对于实 Hilbert 空间 H 中的范数 $\|x\|$ 有

$$\nabla\|x\| = x/\|x\| \quad (0\neq x\in H). \tag{4.1.3}$$

当 $H = \mathbf{R}^n$ 时,公式(4.1.3)在经典分析中是熟知的.

(iii) 仍设 H 是实 Hilbert 空间, $A \in L(H)$ 是对称的,即满足(参照式(3.5.2))

$$\langle Ax, y\rangle = \langle x, Ay\rangle \quad (x,y\in H).$$

设 $f(x) = \langle Ax, x\rangle$, 则当 $h\to 0$ 时,

$$\begin{aligned} f(x+h) - f(x) &= \langle A(x+h), x+h\rangle - \langle Ax, x\rangle \\ &= \langle 2Ax, h\rangle + o(\|h\|), \end{aligned}$$

可见 $\nabla f(x) = 2Ax$. 这可看作微分学公式 $(ax^2)' = 2ax$ 的推广.特别,取 $A = I$ 得 $\nabla\|x\|^2 = 2x (x\in H)$. 这正是熟知的公式 $(x^2)' = 2x$ 的推广.

(iv) 若 $Fx \equiv y_0 \in Y$, 则显然 $F'(x) \equiv 0$; 若 $F\in L(X,Y)$, 则对照式(4.1.1)′与

$$F(x+h) - Fx = Fh \quad (x,h\in X)$$

看出 $F'(x) \equiv F (x\in X)$. 这可看作通常微分公式 $(ax)' = a$ 的推广.应注意不要将 $F'(x) = F$ 误认作 $F'(x) = F(x)$. 实际上应该是

$$F'(x)h = Fh \quad (x,h\in X).$$

对于 Fréchet 微分亦有建立微分规则的问题.通常的微分规则在此都有适当的推广,其中最重要的是**链规则**.

命题 4.1.3(链规则) 设 $F:\Omega\subset X\to Y, G:\Omega'\subset Y\to Z, x\in\Omega, y=Fx\in\Omega'$, $F'(x)$ 与 $G'(y)$ 存在, $H=G\circ F$. 则

$$H'(x)=G'(Fx)F'(x). \tag{4.1.4}$$

证 令 $k=F(x+h)-Fx$, 则 $h\to 0\Rightarrow k\to 0$. $\forall\varepsilon>0$, 取 $\delta>0$, 使得当 $h\in X$, $\|h\|<\delta$ 时,同时有

$$\|k-F'(x)h\|\leqslant\varepsilon\|h\|$$

与

$$\|G(y+k)-Gy-G'(y)k\|\leqslant\varepsilon\|k\|.$$

于是,当 $\|h\|<\delta$ 时,有

$$\begin{aligned}&\|H(x+h)-Hx-G'(y)F'(x)h\|\\=&\|G(y+k)-Gy-G'(y)F'(x)h\|\\\leqslant&\|G(y+k)-Gy-G'(y)k\|+\|G'(y)\|\|k-F'(x)h\|\\\leqslant&\varepsilon(\|k\|+\|G'(y)\|\|h\|)\\\leqslant&\varepsilon(\|F'(x)\|+\varepsilon+\|G'(y)\|)\|h\|.\end{aligned}$$

这对照式(4.1.2),看出式(4.1.4)成立. □

特别,取 $G=T\in L(Y,Z)$ 从(4)得

$$(TF)'(x)=TF'(x). \tag{4.1.5}$$

这意味着:有界线性算子与微分运算可交换(参照式(2.4.9)).

经常用到命题 4.1.3 的以下推论:

$$\frac{\mathrm{d}}{\mathrm{d}t}F(\varphi(t))=F'(\varphi(t))\varphi'(t), \tag{4.1.6}$$

其中 t 是实变量,假定 F 与 φ 均可微.

例 4.1.4 考虑应用上常见的非线性积分算子

$$Fu(x)=\int_a^b f(x,y,u(y))\mathrm{d}y\quad(u\in C(J)), \tag{4.1.7}$$

其中 $J=[a,b](a<b)$, $f(x,y,u)$ 及其偏导数 $f_u(x,y,u)$ 都在 $J\times J\times\mathbf{R}$ 上连续. 对任给 $u\in C(J)$, 显然有 $Fu\in C(J)$. 因此式(4.1.7)定义一映射 $F:C(J)\to C(J)$. 如何求 $F'(u)$? 在 Banach 空间微分学中,并不像在经典微分学中一样有一套系统的求导方法,只能具体问题具体分析. 对于式(4.1.7)表示的 F, 取定 u,h

$\in C(J)$，形式地运用公式(4.1.6)得

$$
\begin{aligned}
F'(u)h &= \frac{\mathrm{d}}{\mathrm{d}t}F(u+th)\bigg|_{t=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\int_a^b f(x,y,u(y)+th(y))\mathrm{d}y\bigg|_{t=0} \\
&= \int_a^b \left[\frac{\mathrm{d}}{\mathrm{d}t}f(x,y,u(y)+th(y))\bigg|_{t=0}\right]\mathrm{d}y \\
&= \int_a^b f_u(x,y,u(y))h(y)\mathrm{d}y.
\end{aligned} \tag{4.1.8}
$$

式(4.1.8)表明，$F'(u)$ 原来就是以 $K(x,y) \triangleq f_u(x,y,u(y))$ 为核的线性积分算子(暂时记为 T). 因 $K(x,y)$ 在 $J\times J$ 上连续，故由命题 2.2.4 有 $T \in L(C(J))$. 式(4.1.8)的严格证明如下. 固定 $u\in C(J)$，作以下估计：

$$
\begin{aligned}
&\| F(u+h) - Fu - Th \|_0 \\
&= \sup_{x\in J}\left|\int_a^b [f(x,y,u(y)+h(y)) - f(x,y,u(y)) - f_u(x,y,u(y))h(y)]\mathrm{d}y\right| \\
&= \sup_{x\in J}\left|\int_a^b [f_u(x,y,u(y)+\theta h(y)) - f_u(x,y,u(y))]h(y)\mathrm{d}y\right| \\
&\qquad\qquad (0<\theta=\theta(y)<1,\text{用 Lagrange 中值公式}) \\
&\leqslant \sup_{x\in J}\int_a^b | f_u(x,y,u(y)+\theta h(y)) - f_u(x,y,u(y)) | \mathrm{d}y \| h \|_0 \\
&= o(\| h \|_0),\quad (\| h \|_0 \to 0)
\end{aligned}
$$

最后一步用到 $f_u(x,y,u)$ 的连续性. 将以上结果与(1)对照看出 $F'(u) = T$，即式(4.1.8)成立.

在条件足够强的情况下，通常采用一个如导出式(4.1.8)那样的简化程序，而略去较繁琐的严格论证.

往下要做的事如同经典微分学中一样：建立微分中值定理、Taylor 公式等. 但你立即遇到一个障碍：Lagrange 中值定理不适用于向量值函数，即使在有限维空间中也是如此. 例如，复函数(可看作 2 维向量值函数)

$$z(t) = \mathrm{e}^{\mathrm{i}t}\quad (t\in\mathbf{R})$$

尽管满足 $z(0) = z(2\pi) = 1$，但处处有 $z'(t) = \mathrm{i}z(t) \neq 0$. 因此，推广后的中值定理不是取"Lagrange 中值"形式，而是取不等式形式. 不过，这并没有什么不好，因

在很多情况下,两者的作用是一样的.

定理 4.1.5(中值定理) 设 $F:\Omega\subset X\to Y$,线段

$$[x,y]\triangleq\{(1-t)x+ty:0\leqslant t\leqslant 1\}$$

含于 Ω 内.若 F 在 $[x,y]$ 上每点可微,则成立

$$\|Fx-Fy\|\leqslant\sup_{z\in[x,y]}\|F'(z)\|\,\|x-y\|. \tag{4.1.9}$$

证 以 β 记式(4.1.9)之右端,只需证(用式(2.4.14))

$$|\langle v,Fx-Fy\rangle|\leqslant\beta\|v\|\quad(\forall v\in Y^*).$$

取定 $v\in Y^*$,令 $\varphi(t)=\langle v,F((1-t)x+ty)\rangle$.由式(4.1.5),(4.1.6)有

$$\varphi'(t)=\langle v,F'(z_t)(y-x)\rangle,\quad z_t=(1-t)x+ty.$$

于是,用通常的 Lagrange 中值定理得:

$$\begin{aligned}|\langle v,Fx-Fy\rangle|&=|\varphi(1)-\varphi(0)|=|\varphi'(\tau)|\quad(\tau\in(0,1))\\&=|\langle v,F'(z_\tau)(y-x)\rangle|\\&\leqslant\|v\|\,\|F'(z_\tau)\|\,\|y-x\|\leqslant\beta\|v\|,\end{aligned}$$

即如所要证. □

若以 $F-T$ 代 F,$T\in L(X,Y)$,则从(4.1.9)得到:

$$\|Fx-Fy-T(x-y)\|\leqslant\sup_{z\in[x,y]}\|F'(z)-T\|\,\|x-y\|. \tag{4.1.10}$$

为建立 Taylor 公式,还需作一点准备.关键是如何解释 Taylor 公式中出现的 $F^{(n)}(x)h^n$.设 $F:\Omega\subset X\to Y$ 在 $x\in\Omega$ 存在 n 阶导数,则 $F^{(n)}(x)$ 诱导出一个 n 元映射

$$T:X^n\triangleq\overbrace{X\times X\times\cdots\times X}^{n\text{ 个}}\to Y$$

$$(h_1,h_2,\cdots,h_n)\to(\cdots(F^{(n)}(x)h_1)h_2\cdots)h_n.$$

显然 $T(h_1,h_2,\cdots,h_n)$ 对每个 $h_i(1\leqslant i\leqslant n)$ 是线性的,且

$$\|T(h_1,h_2,\cdots,h_n)\|\leqslant\|F^{(n)}(x)\|\,\|h_1\|\,\|h_2\|\cdots\|h_n\|. \tag{4.1.11}$$

称这样的 T 为有界 n 重线性算子,仍将 T 记作 $F^{(n)}(x)$,且约定

$$F^{(n)}(x)h_1h_2\cdots h_n = T(h_1,h_2,\cdots,h_n);$$

$$F^{(n)}(x)h^n = T(h,h,\cdots,h).$$

当然,此处 h^n 不应看作 h 的 n 次方,除非 $\dim X = 1$. 不过, $F^{(n)}(x)h^n$ 确类似于 n 次幂函数. 例如, $\forall t \in \mathbf{R}$, 有 $F^{(n)}(x)(th)^n = t^nF^{(n)}(x)h^n$; 再如,由不等式(4.1.11)推出

$$\| F^{(n)}(x)h^n \| \leqslant \| F^{(n)}(x) \| \| h \|^n. \tag{4.1.12}$$

注意 $F^{(n)}(x)h^n$ 正是与通常的 n 阶微分 $F^{(n)}(x)\mathrm{d}x^n$ 相当的东西.

定理 4.1.6 (Taylor 公式)　设 $F:\Omega\subset X\to Y$ 是 C^{n+1} 映射, $0\leqslant n<\infty$, Y 完备,线段 $[x,x+h]$ 含于 Ω. 则成立如下 **Taylor 公式**:

$$F(x+h) = \sum_{k=0}^{n} \frac{1}{k!}F^{(k)}(x)h^k + \frac{1}{n!}\int_0^1 (1-t)^nF^{(n+1)}(x+th)h^{n+1}\mathrm{d}t.$$

$$(F^{(0)}(x)h^0 = F(x)) \tag{4.1.13}$$

证　因 $F^{(n+1)}(\cdot)$ 连续, Y 完备,故式(4.1.13)中的积分存在(参考§2.4.3). 当 $n=0$ 时,式(4.1.13)成为

$$F(x+h) - Fx = \int_0^1 F'(x+th)h\mathrm{d}t = \int_0^1 \left[\frac{\mathrm{d}}{\mathrm{d}t}F(x+th)\right]\mathrm{d}t,$$

这就是 Newton-Leibniz 公式(2.4.8). 若 $n\geqslant 1$, 则依归纳法有

$$F(x+h) - \sum_{k=0}^{n-1} \frac{1}{k!}F^{(k)}(x)h^k$$

$$= \frac{1}{(n-1)!}\int_0^1 (1-t)^{n-1}F^{(n)}(x+th)h^n\mathrm{d}t \qquad \text{(用归纳假设)}$$

$$= \frac{1}{n!}(1-t)^nF^{(n)}(x+th)h^n\Big|_1^0 + \frac{1}{n!}\int_0^1 (1-t)^n\left[\frac{\mathrm{d}}{\mathrm{d}t}F^{(n)}(x+th)h^n\right]\mathrm{d}t \qquad \text{(分部积分)}$$

$$= \frac{1}{n!}F^{(n)}(x)h^n + \frac{1}{n!}\int_0^1 (1-t)^n F^{(n+1)}(x+th)h^{n+1}\mathrm{d}t,$$

这正好得出要证的等式(4.1.13). □

若以 $R_n(h)$ 记式(4.1.13)中的**积分余项**,令 $M = \sup\limits_{0\leqslant t\leqslant 1}\|F^{(n+1)}(x+th)\|$,则

$$\|R_n(h)\| \leqslant \frac{1}{n!}\int_0^1 (1-t)^n M\|h\|^{n+1}\mathrm{d}t \quad (\text{用式}(4.1.12))$$

$$= \frac{M\|h\|^{n+1}}{(n+1)!}.$$

因此,当 $\|h\|\to 0$ 时从式(4.1.13)得出

$$\left\|F(x+h) - \sum_{k=0}^{n}\frac{1}{k!}F^{(k)}(x)h^k\right\| = o(\|h\|^n).$$

§4.1.2 偏导数

如果你顺着经典微分学的思路考虑,在有了上面的结果之后,自然会认为该进入"多元函数微分学"了.不过,现在必须澄清一点,在抽象空间的框架内,概念上并无"一元函数"与"多元函数"之分.一个 $W = X\times Y\times Z$ 上的"三元函数" $G(x,y,z)$ 既可看作 W 上的一元函数 $G(w)$, $w=(x,y,z)$, 又可看作 $X\times(Y\times Z)$ 上的二元函数 $G(x,v)$, $v=(y,z)$. 不过,从方法上考虑,分别研究 $G(x,y,z)$ 对各变元的依赖性通常是有利的,这就使偏导数概念仍有其必要.偏导数就是对于各变元的导数,其定义如同经典微分学中一样;自然沿用通常的偏导数记号,如

$$F_x, F_x(x,y), \quad \frac{\partial F(x,y)}{\partial x}, \quad \frac{\partial^2 F(x,y)}{\partial x\partial y}.$$

惟一需要强调的是,对于 $F:\Omega\subset X\times Y\to Z$, 当其偏导数存在时有

$$F_x(x,y)\in L(X,Z), \quad F_y(x,y)\in L(Y,Z),$$

这两个偏导数是不同空间上的线性算子(除非 $X = Y$),互相之间并无直接联系.对于高阶偏导数亦有类似情况.

若令 $z=(x,y)$, 自然发生问题:三个导数

$$F'(z), \quad F_x(x,y), \quad F_y(x,y)$$

之间关系如何?这正是如下定理所要回答的.

定理 4.1.7 设 $X_j, Y_i(1\leqslant j\leqslant n, 1\leqslant i\leqslant m)$ 是赋范空间，$X=\prod_j X_j, Y=\prod_i Y_i, \Omega\subset X$ 是一非空开集，给定映射

$$\begin{cases}F=(F_1,F_2,\cdots,F_m):\Omega\subset X\to Y,\\ y_i=F_i(x), x=(x_1,\cdots,x_n)\in\Omega, y=(y_1,\cdots,y_m)\in Y.\end{cases}$$

则以下结论成立：

(i) $F\in C^r(1\leqslant r<\infty)\Leftrightarrow$ 每个 F_i 存在所有 r 阶连续偏导数.

(ii) 若 $F'(x)$ 存在，则 $\partial F_i(x)/\partial x_j$ 均存在，且

$$F'(x)h=\left(\sum_j\frac{\partial F_1(x)}{\partial x_j}h_j,\cdots,\sum_j\frac{\partial F_m(x)}{\partial x_j}h_j\right), \tag{4.1.14}$$

其中 $h=(h_1,h_2,\cdots,h_n)\in X$. 若将式(4.1.14)右端与 h 都写成列，则式(4.1.14)可改写成：

$$F'(x)h=\begin{bmatrix}\dfrac{\partial F_1(x)}{\partial x_1} & \dfrac{\partial F_1(x)}{\partial x_2} & \cdots & \dfrac{\partial F_1(x)}{\partial x_n}\\ \dfrac{\partial F_2(x)}{\partial x_1} & \dfrac{\partial F_2(x)}{\partial x_2} & \cdots & \dfrac{\partial F_2(x)}{\partial x_n}\\ \vdots & \vdots & & \vdots\\ \dfrac{\partial F_m(x)}{\partial x_1} & \dfrac{\partial F_m(x)}{\partial x_2} & \cdots & \dfrac{\partial F_m(x)}{\partial x_n}\end{bmatrix}\begin{bmatrix}h_1\\ h_2\\ \vdots\\ h_n\end{bmatrix} \tag{4.1.14$'$}$$

证 以 $P_j:X\to X_j$ 与 $Q_i:Y\to Y_i$ 记投影，以 $J_j:X_j\to X$ 与 $K_i:Y_i\to Y$ 记嵌入，例如 $J_1x_1=(x_1,0,\cdots,0)$. $P_j,Q_i,J_j,K_i(1\leqslant j\leqslant n,1\leqslant i\leqslant m)$ 均为有界线性算子. 显然

$$F=\sum_i K_iF_i,\quad F_i=Q_iF(1\leqslant i\leqslant m). \tag{4.1.15}$$

首先设 $F'(x)$ 存在. 因映射 $x_j\to F_i(x_1,\cdots,x_j,\cdots,x_n)$ 可分解成：

$$x_j\xrightarrow{J_j}(0,\cdots,x_j,\cdots,0)\xrightarrow{\text{平移}}(x_1,\cdots,x_j,\cdots,x_n)\triangleq x$$
$$\xrightarrow{F}F(x)\xrightarrow{Q_i}F_i(x),$$

于是由链规则与公式(4.1.5)有

$$\frac{\partial F_i(x)}{\partial x_j}=Q_iF'(x)J_j\quad(1\leqslant j\leqslant n,1\leqslant i\leqslant m). \tag{4.1.16}$$

任给 $h=(h_1,h_2,\cdots,h_n)\in X$，有 $h=\sum J_j h_j$，于是由式(4.1.16)有

$$\sum_j \frac{\partial F_i(x)}{\partial x_j}h_j = Q_i F'(x)\sum_j J_j h_j = Q_i F'(x)h,$$

这正表明式(4.1.14)成立.因此结论(ii)得证.

为证结论(i)，只需考虑 $r=1$ 的情况(一般情况由归纳法得出).若 $F\in C^1$，则由已证结论知式(4.1.16)成立，且式(4.1.16)右端显然是 x 的连续函数.反之，设所有偏导数 $\partial F_i/\partial x_j$ 存在且在 Ω 内连续，今证 $F\in C^1$. 由式(4.1.15)，只需证 $F_i\in C^1(1\leqslant i\leqslant m)$；因此不妨设 $m=1, F=F_1$. 若能证

$$F'(x)=\sum_j \frac{\partial F(x)}{\partial x_j}P_j, \tag{4.1.17}$$

则 $F'(x)$ 显然对 x 连续，因而 $F\in C^1$. 式(4.1.17)相当于

$$F'(x)h=\sum_j \frac{\partial F(x)}{\partial x_j}h_j \quad (h=(h_j)\in X). \tag{4.1.17$'$}$$

今证式(4.1.17)′.为简便起见，不妨设 $n=2$. 式(4.1.17)′由以下估计得出：

$$\begin{aligned}
&\left\| F(x+h)-F(x)-\sum_j \frac{\partial F(x)}{\partial x_j}h_j \right\| \\
\leqslant &\left\| F(x+h)-F(x_1,x_2+h)-\frac{\partial F(x)}{\partial x_1}h_1 \right\| \\
&+\left\| F(x_1,x_2+h)-F(x)-\frac{\partial F(x)}{\partial x_2}h_2 \right\| \\
\leqslant &\sup_{0\leqslant t\leqslant 1}\left\| \frac{\partial F(x_1+th_1,x_2+h)}{\partial x_1}-\frac{\partial F(x)}{\partial x_1} \right\| \| h_1\| \\
&+\sup_{0\leqslant t\leqslant 1}\left\| \frac{\partial F(x_1,x_2+th)}{\partial x_2}-\frac{\partial F(x)}{\partial x_2} \right\| \| h_2\| \qquad \text{(用式(4.1.10))} \\
=&\, o(\| h_1\|+\| h_2\|). \qquad \text{(用 } \partial F/\partial x_j \text{ 的连续性)}
\end{aligned}$$

□

容易看出，式(4.1.17)′可看作通常的全微分公式的推广；而 $F'(x)$ 不妨称为**全导数**. 式(4.1.14)与式(4.1.17)恰好表达了全导数与偏导数的关系. 式(4.1.14)′中的算子矩阵

$$J(x)=\left[\frac{\partial F_i(x)}{\partial x_j}\right]_{m\times n}$$

称为 F 在 x 的 **Jacobi 矩阵**. 这样, 可将式(4.1.14)$'$缩写成

$$F'(x)h = J(x)h \quad (h \in X).$$

通常就认定 $F'(x) = J(x)$. 在 $m = 1$ 与 $n = 1$ 这两种情况下分别有:

$$F'(x) = \left[\frac{\partial F(x)}{\partial x_1}, \frac{\partial F(x)}{\partial x_2}, \cdots, \frac{\partial F(x)}{\partial x_n}\right]$$

与

$$F'(x) = [F'_1(x), F'_2(x), \cdots, F'_m(x)]^{\mathrm{T}}.$$

特别, 若取 $X = \mathbf{R}^n$, $Y = \mathbf{R}^m$, 则当 $F: \Omega \subset \mathbf{R}^n \to \mathbf{R}^m$ 可微时, $F'(x)$ 就是通常的 Jacobi 矩阵: $F'(x) = [\partial F_i(x)/\partial x_j]$. 特别, 当 $X = \mathbf{R}^n$, $Y = \mathbf{R}$, $f: \Omega \subset \mathbf{R}^n \to \mathbf{R}$ 可微时, 有

$$\nabla f(x) = \left[\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \cdots, \frac{\partial f(x)}{\partial x_n}\right],$$

这与通常的梯度概念一致.

§4.1.3　G 导数

至此你已看到, 基于 Fréchet 导数的微分学以相当完满的形式推广了经典微分学. 从前面所展开的内容来看, 你很难察觉到空间的无限维性带来了什么影响. 似乎能得出结论: 对于微分理论来说, 空间的维数并不是一个重要因素. 如果仅从微分公式的形式方面考虑, 事情确实如此. 但若考虑到可微性, 有限维空间与无限维空间就完全不可同日而语了. 你只要想想: 无限维空间上的线性函数也可能是处处不可微的. 要害问题是: 在无限维空间中, 定义 4.1.1 意义下的可微性是一个极强的条件. 如果很少有映射是可微的, 那么所建立的微分理论将是一个形式优美的**空洞理论**. 当然, 事情并未严重到应完全摈弃 Fréchet 微分理论的地步; 但它不能完全满足应用之需要则是肯定的. 这就有必要考虑某些条件较弱的微分概念, 所谓 G 导数正是其中较常用的一种.

定义 4.1.8　设 $F: D \subset X \to Y$, $x \in D$, $h \in X$. 称

$$\mathrm{D}_+ F(x, h) = \lim_{t \downarrow 0} \frac{F(x + th) - Fx}{t} \tag{4.1.18}$$

为 F 在 x 沿 h 的**右侧方向导数**, 只要上式右端之极限存在. 类似地, 称

$$\mathrm{D}_- F(x, h) = \lim_{t \uparrow 0} \frac{F(x + th) - Fx}{t} \tag{4.1.18}$$

为 F 在 x 沿 h 的**左侧方向导数**. 当以上两个方向导数都存在且相等时, 就称其为 F

在 x 沿 h 的**方向导数**,并记作 $\mathrm{D}F(x,h)$. 若 $\forall h \in X$, $\mathrm{D}F(x,h)$ 存在,且 $\mathrm{D}F(x,\cdot) \in L(X,Y)$,则说 F 在 x 处**G 可微**,并称 $\mathrm{D}F(x) \triangleq \mathrm{D}F(x,\cdot)$ 为 F 在 x 的**G 导数**[①].

对于 $f: D \subset X \to \mathbf{R}$, 若 G 导数 $\mathrm{D}f(x)$ 存在,通常也称 $\mathrm{D}f(x)$ 为 $f(x)$ 的梯度,并记作 $\nabla f(x)$. 若方向导数 $\mathrm{D}f(x,h)$ 存在,则在传统上称它为 f 在 x 关于增量 h 的**一阶变分**,也记作 $\delta f(x,h)$. 变分概念源于**变分学**,即研究泛函极值的一个数学分支,其历史可追溯到微积分学的开创时期. 普遍认为,变分学是泛函分析的渊源之一.

想必你希望立即知道: G 导数与 F 导数的关系如何? 这由以下命题解答.

命题 4.1.9 设 $F: \Omega \subset X \to Y$, $x \in \Omega$. 则 F 导数 $F'(x)$ 存在的充要条件是: G 导数 $\mathrm{D}F(x)$ 存在,且关于 $h \in X$, $\|h\| = 1$ 一致地有

$$\mathrm{D}F(x,h) = \lim_{t \to 0} \frac{F(x+th) - Fx}{t}. \tag{4.1.19}$$

当 $F'(x)$ 存在时必 $\mathrm{D}F(x) = F'(x)$, $\mathrm{D}F(x,h) = F'(x)h$.

证明是容易的,你自己不难写出它. 注意(4.1.19)可写作

$$\mathrm{D}F(x,h) = \frac{\mathrm{d}}{\mathrm{d}t}F(x+th)\Big|_{t=0}. \tag{4.1.19$'$}$$

关于方向导数还需作两点解释. 首先,由式(4.1.19)看出,在任何情况下总有 $\mathrm{D}F(x,0) = 0$, 故只有沿"非零方向" h 的方向导数才真正有意义. 其次,由式(4.1.18)看出, $\mathrm{D}_{\pm} F(x,h)$ 对 h 是正齐次的,即

$$\mathrm{D}_{\pm} F(x,\alpha h) = \alpha \mathrm{D}_{\pm} F(x,h) \quad (\alpha > 0),$$

只要上式右端存在. 因此,只要考虑对单位向量 h (即 $\|h\| = 1$)的方向导数就够了. 在经典微分学中,实际上就是在这个意义上使用方向导数. 还可指出左侧与右侧方向导数之间的以下简单联系:

$$\mathrm{D}_{+} f(x,-h) = -\mathrm{D}_{-} f(x,h). \tag{4.1.20}$$

要解释以上概念,用有限维例子也很能说明问题.

例 4.1.10 (i) 设 $f = \chi_{[0,\infty)}$, 则

$$\mathrm{D}_{+} f(0,1) = \lim_{t \downarrow 0} \frac{f(t) - f(0)}{t} = 0;$$

---

① 定义 4.1.8 中所给出的概念,早已见于文献,只是所用术语与记号颇不统一. 读者在其他处看到类似概念时,务必首先确定其内涵,以免混淆.

而 $\mathrm{D}_- f(0,1)$ 不存在(实际上, $\mathrm{D}_- f(0,1) = \infty$).

(ii) 设 $f(x)=|x|$ $(x\in\mathbf{R})$, 则

$$\lim_{t\downarrow 0}\frac{f(0\pm t)-f(0)}{\pm t}=\pm 1,$$

可见 $\mathrm{D}_\pm f(0,1)=\pm 1$: 两侧方向导数存在,但不相等.

(iii) 设 $f(x)=x_1x_2^2/|x|^2(0\neq x=(x_1,x_2)\in\mathbf{R}^2)$, $f(0)=0$. 则对任给 $0\neq h\in\mathbf{R}^2$, 有

$$\mathrm{D}f(0,h)=\lim_{t\to 0}\frac{f(th)-f(0)}{t}=f(h).$$

可见 f 在 $x=0$ 沿任何方向 h 的方向导数存在,但 $\mathrm{D}f(0,h)$ 不是 h 的线性函数,于是 G 导数 $\mathrm{D}f(0)$ 不存在,因而 F 导数 $f'(0)$ 必不存在(命题 4.1.9).

§4.2 压缩映射与迭代法

判定方程的可解性及计算方程的近似解,始终是数学及其应用中的基本课题. 本节给出解决这类问题的一个普遍方法.

§4.2.1 压缩映射原理

给定一个映射 $F: D\subset X\to Y$. 若 $D=X$ 且 F 是线性的,则依式(2.1.5)有

$$\|F\|=\sup_{x\neq y}\frac{\|Fx-Fy\|}{\|x-y\|}. \tag{4.2.1}$$

算子范数作为量度线性算子的“伸张系数”的数量指标(参看 §2.1.2),已显示出极大的重要性. 现在让我们试探一下,类似的概念能否在非线性算子的研究中发挥作用. 实际上,无论 F 是否为线性算子,只要式(1)之右端(记为 β)有限,我们就有估计式

$$\|Fx-Fy\|\leqslant\beta\|x-y\|,\quad (x,y\in D)$$

而且这个估计一般地已不能再改进. 如上的估计在实际应用中的重要性,是不言而喻的. 这就表明,对于非线性算子亦需要有一个类似于算子范数的概念. 它就是如下给出的:

定义 4.2.1 设 $F: D\subset X\to Y$, 令

$$\mathrm{Lip}F=\sup_{x,y\in D,x\neq y}\frac{\|Fx-Fy\|}{\|x-y\|}. \tag{4.2.2}$$

若 $\mathrm{Lip}F<\infty$, 则称 F 为从 D 到 Y 的 **Lipschitz 映射**,称 $\mathrm{Lip}F$ 为 F 的 **Lipschitz 模**

数.若 $\mathrm{Lip}F \leqslant 1$, 则称 F 为非扩张映射;若 $\mathrm{Lip}F < 1$, 则称 F 为**压缩映射**.

对照式(4.2.1)与式(4.2.2)看出,若 F 是线性的且 $D = X$, 则 $\mathrm{Lip}F = \| F \|$. 可见,**Lipschitz 模数**正是算子范数的"非线性推广". 鉴于此,你很容易将 §2.1 中涉及算子范数的某些结论推广于 Lipschitz 模数. 例如,

$$\| Fx - Fy \| \leqslant (\mathrm{Lip}F) \| x - y \| \quad (x, y \in D) \tag{4.2.3}$$

(对照式(2.1.6)),式(4.2.3)就是对 $\| Fx - Fy \|$ 的某种"最佳估计",

$$\| Fx - Fy \| \leqslant \beta \| x - y \| \ (\forall x, y \in D) \Rightarrow \mathrm{Lip}F \leqslant \beta, \tag{4.2.4}$$

$$\begin{cases} \mathrm{Lip}(\alpha F) = | \alpha | \mathrm{Lip}F, \\ \mathrm{Lip}(F + \mathrm{const}) = \mathrm{Lip}F, \\ \mathrm{Lip}(F_1 + F_2) \leqslant \mathrm{Lip}F_1 + \mathrm{Lip}F_2, \\ \mathrm{Lip}(G \circ F) \leqslant \mathrm{Lip}G \ \mathrm{Lip}F. \end{cases} \tag{4.2.5}$$

Lipschitz 映射显然是连续的(参照命题 2.1.4),但逆命题不再成立.

在具体情况下,很少能准确地求出 $\mathrm{Lip}F$, 这对于 Lipschitz 模数的应用并不构成本质的限制.通常若能得出某一估计 $\mathrm{Lip}F \leqslant \beta$, 使得 β 能满足问题的需要(例如 $\beta < 1$)就够了.这与对线性算子范数的要求是类似的.对于可微映射 F, 可用中值定理来估计 $\mathrm{Lip}F$: 若 $F: \Omega \subset X \to Y$ 是可微映射, Ω 是凸开集,则由式(4.1.9)有

$$\| Fx - Fy \| \leqslant \sup_{z \in \Omega} \| F'(z) \| \, \| x - y \|,$$

因此由式(4.2.4)有

$$\mathrm{Lip}F \leqslant \sup_{z \in \Omega} \| F'(z) \|. \tag{4.2.6}$$

由式(4.2.6)推出,若在 Ω 内 $\| F'(x) \|$ 适当小,则 F 是压缩映射.这是判定压缩映射的最常用的方法之一.

例 4.2.2 (i) 设 $F: \Omega \subset \mathbf{R}^n \to \mathbf{R}^n$ 为 C^1 映射, $x_0 \in \Omega$, $A = F'(x_0)$ 是对称矩阵(参看定理 4.1.7 之后的说明), $\sigma(A) \subset (-1, 1)$. 由定理 3.4.5(ii),从 $\sigma(A) \subset (-1, 1)$ 得出 $\| A \|_2 < 1$. 由 $F'(x)$ 对 x 的连续性,必有某个 $r > 0$, 使得在球 $\bar{B}_r(x_0)$ 内 $\| F'(x) \|_2 < 1$, 因而

$$\beta \triangleq \max_{|x - x_0| \leqslant r} \| F'(x) \|_2 < 1.$$

于是由式(4.2.6)得出: F 在球 $\bar{B}_r(x_0)$ 上为压缩映射.

(ii) 设 $J = [a, b] (a < b)$, $f \in C(J \times J \times \mathbf{R}, \mathbf{R})$,

$$|f_z(x,y,z)| \leqslant \beta. \quad ((x,y,z) \in J \times J \times \mathbf{R})$$

令

$$Fu(x) = \int_a^x f(x,y,u(y))\mathrm{d}y, \quad (x \in J, u \in C(J)) \tag{4.2.7}$$

则得到一个非线性算子 $F: C(J) \to C(J)$. $\forall u, v \in C(J)$, 有

$$\begin{aligned} \| Fu - Fv \|_0 &= \sup_{x \in J} \left| \int_a^x [f(x,y,u(y)) - f(x,y,v(y))]\mathrm{d}y \right| \\ &\leqslant \sup_{x \in J} \int_a^b |f(x,y,u(y)) - f(x,y,v(y))| \mathrm{d}y \\ &\leqslant \int_a^b \beta |u(y) - v(y)| \mathrm{d}y \qquad \text{(用中值定理)} \\ &\leqslant \beta(b-a) \| u - v \|_0, \end{aligned}$$

由此得 $\mathrm{Lip}F \leqslant \beta(b-a)$. 若 $\beta(b-a) < 1$, 则 F 是压缩映射.

对于 Lipschitz 映射有许多精细的研究. 但本节的主要对象是压缩映射. 压缩映射的特殊价值在于,它对于判定方程可解及解的近似计算起关键作用. 下面就来说明如何做到这一点.

在代数方程、微分方程、积分方程、泛函方程等诸多问题的研究中,人们逐渐形成了一种共识:如果将所考虑的方程问题转化为所谓**不动点方程**

$$x = Fx \tag{4.2.8}$$

(F 是某一算子,一般是非线性的;方程(4.2.8)的解称为 F 的**不动点**),从而将求解方程的问题变成求 F 的不动点的问题,那么就可以充分利用(线性与非线性)算子理论中的丰富成果,尤其是利用各种不动点定理. 在现代数学的应用中,不动点定理的重大价值与普遍使用是众所周知的,其中最简单且最常用的一个就是如下原理.

定理 4.2.3(压缩映射原理)[①]　设 $D \subset X$ 是一非空闭集, $F: D \to D$.

(i) 不动点的存在惟一性. 若对某个 $n \geqslant 1$ 有 $\mathrm{Lip}F^n < 1$, 则 F 在 D 上有惟一不动点 x^*.

(ii) 迭代序列. 设 $\mathrm{Lip}F < 1$, 任给 $x_0 \in D$, 令

$$x_n = Fx_{n-1} = F^n x_0 \quad (n \geqslant 1), \tag{4.2.9}$$

① 又称为 Banach 不动点定理,它实际上亦适用于度量空间中的压缩映射. 不过,其主要应用无疑是在 Banach 空间中考虑,因此拟不追求更大的一般性.

则如此得到的**迭代序列** $\{x_n\}$ 收敛于 F 在 D 上的惟一不动点 x^*.

(iii) 误差估计. 设 $\mathrm{Lip}F \leqslant r < 1$, x_n 依式(4.2.9), $x^* \in D$ 是 F 的不动点,则

$$\| x_n - x^* \| < \frac{r^n}{1-r} \| x_0 - Fx_0 \| ; \quad (\text{先验估计}) \qquad (4.2.10)$$

$$\| x_n - x^* \| < \frac{r}{1-r} \| x_n - x_{n-1} \| . \quad (\text{后验估计}) \qquad (4.2.11)$$

证 设 $\mathrm{Lip}F \leqslant r < 1$. 任取 $x_0 \in D$, 设 x_n 依式(4.2.9). 则

$$\| x_{n+1} - x_n \| = \| F^n x_1 - F^n x_0 \| \leqslant (\mathrm{Lip}F^n) \| x_1 - x_0 \| \quad (\text{用式}(4.2.3))$$

$$\leqslant (\mathrm{Lip}F)^n \| x_1 - x_0 \| \leqslant r^n \| x_1 - x_0 \| ; \quad (\text{用式}(4.2.5))$$

$$\| x_m - x_n \| \leqslant \sum_{i=n}^{m-1} \| x_{i+1} - x_i \| \leqslant \sum_{i=n}^{m-1} r^i \| x_1 - x_0 \|$$

$$\leqslant \frac{r^n}{1-r} \| x_1 - x_0 \| . \quad (m > n) \qquad (4.2.12)$$

式(4.2.12)表明 $\{x_n\}$ 是 Cauchy 列. 因 X 完备而 D 是闭集,故 $x_n \to x^* \in D$. 在等式 $x_n = Fx_{n-1}$ 两边令 $n \to \infty$ 得 $x^* = Fx^*$, 可见 x^* 是 F 的不动点. 若此外 F 还有不动点 $x \in D$, 则

$$\| x - x^* \| = \| Fx - Fx^* \| \leqslant r \| x - x^* \| .$$

因 $r < 1$, 上式推出 $\| x - x^* \| = 0$, 从而 $x = x^*$. 可见 x^* 是 F 在 D 上的惟一不动点. 这就证明了结论(ii).

其次,在不等式(4.2.12)中令 $m \to \infty$, 即得误差估计(4.2.10);然后将式(4.2.10)用到迭代序列 $\{x_{n-1}, x_n, x_{n+1}, \cdots\}$ (这相当于以 x_{n-1} 作为新的初始点)且只迭代一次,即得式(4.2.11). 因此(iii)得证.

若 $\mathrm{Lip}F^m < 1$, 则 F^m 在 D 上有惟一不动点 $\bar{x}$. 因

$$F^m(F\bar{x}) = F(F^m\bar{x}) = F\bar{x},$$

故 $F\bar{x}$ 亦是 F^m 在 D 上的不动点, 于是由惟一性有 $F\bar{x} = \bar{x}$, 即 $\bar{x}$ 是 F 在 D 上的不动点. 若 $x \in D$ 亦是 F 的不动点,则 x 显然也是 F^m 的不动点,因而必定 $x = \bar{x}$. 因此 $\bar{x}$ 是 F 在 D 上的惟一不动点. 这就证得(i). 于是定理证毕. □

压缩映射原理是现代数学中极令人惊异的典型例子:其所依据的思想及证明都至为简单,而其应用则异常广泛且富有成效. 下面考虑压缩映射原理的两类典型应用.

§4.2.2　用于存在性证明

压缩映射原理首先用来证明各种类型方程的解的存在性(通常也同时得出惟一性),这类应用涉及函数方程、微分方程、积分方程等范围广泛的问题.解决这类问题的通常步骤是:

(i) 构成空间 X,它应包含所考虑方程的解.

(ii) 转化方程,即将所考虑的方程依适当方式转化为不动点方程(4.2.8),这就自然得到一个算子 $F: D(F) \subset X \to X$.

(iii) 选取非空闭集 $D \subset D(F)$,使得 $FD \subset D$. 有时这是一个困难的选择:D 取得过小,可能使 $FD \subset D$ 不满足;D 取得过大,可能使映射的压缩性难以验证.

(iv) 验证 $F: D \to D$ 或某个 F^m 为压缩映射,这通常是整个论证的核心部分.

(v) 应用压缩映射原理.在作了(i)~(iv)之后,这当然是平凡的一步.但也应注意,由原理4.2.3仅仅得出:原方程在 D 中仅有一解;至于它是否有集 D 以外的解,则是另一个问题.

现在考虑一些具体例子.

例 4.2.4(微分方程存在定理)　考虑 Banach 空间 X 中的微分方程初值问题:

$$y'(t) = f(t, y(t)), \quad y(a) = x_0 \in X, \tag{4.2.13}$$

其中 $f \in C(J \times X, X)$, $J = [a, b]$ $(a < b)$. 若 f 满足 Lipschitz 条件:

$$\| f(t,x) - f(t,y) \| \leqslant \beta \| x - y \| \quad (t \in J, x, y \in X), \tag{4.2.14}$$

则问题(4.2.13)式在 J 上有惟一解.

证　(i) 构成空间.令 $Y = C(J, X)$,在 Y 中定义范数:

$$\| y \| = \sup_{t \in J} \| y(t) \| \quad (y \in Y).$$

验证如上的范数满足范数公理,且 Y 依此范数完备,是一件平凡的事情.不过,若你能当作练习来做将不无益处.

(ii) 转化方程.如微分方程论中通常所作的,问题(4.2.13)式可转化为积分方程:

$$y(t) = x_0 + \int_a^t f(s, y(s)) \mathrm{d}s \triangleq Fy(t). \tag{4.2.15}$$

这已经是一个不动点方程:$y = Fy$;算子 F 已由(4.2.15)式同时得到定义.

(iii) 直接看出,对任给 $y \in Y$,有 $Fy \in Y$,故就以 Y 作为定理4.2.3中的闭集 D.

(iv) 验证 F^n (n 充分大)为压缩映射,这是整个论证的关键.利用 F 的定义不难归纳得出:

$$F^n y(t) = \sum_{k=0}^{n-1} \frac{(t-a)^k}{k!} x_0 + \int_a^t \mathrm{d}t_1 \int_a^{t_1} \mathrm{d}t_2 \cdots \int_a^{t_{n-1}} f(t_n, y(t_n)) \mathrm{d}t_n,$$

其中 $y \in Y$. 于是对任给 $y, z \in Y$ 有

$$\begin{aligned} \| F^n y - F^n z \| &\leqslant \sup_{t \in J} \int_a^t \mathrm{d}t_1 \int_a^{t_1} \mathrm{d}t_2 \cdots \int_a^{t_{n-1}} | f(s, y(s)) - f(s, z(s)) | \mathrm{d}s \\ &\leqslant \int_a^b \mathrm{d}t_1 \int_a^{t_1} \mathrm{d}t_2 \cdots \int_a^{t_{n-1}} \beta \| y - z \| \mathrm{d}s \qquad (\text{用式}(4.2.14)) \\ &= \frac{\beta(b-a)^n}{n!} \| y - z \|, \end{aligned}$$

因而 $\mathrm{Lip} F^n \leqslant \beta(b-a)^n / n! < 1$ (n 充分大)①.

余下的事情只是应用定理 4.2.3 了. □

特别取 $X = \mathbf{R}^n$, 就从例 4.2.4 得出在通常微分方程教材中已熟知的存在惟一性定理.略加对照之后你必定会看出,运用泛函分析的语言、结论与方法,对于这一熟知定理的证明被大大简化了.

在例 4.2.4 的证明中,我们严格地按照标准的步骤进行表述,以使读者获得尽可能清晰的印象.不过,你一旦熟练地掌握了这一方法,在实际运用时就完全可以变通处理,而不必拘泥于某一固定格式.

例 4.2.5(Volterra 型积分方程) 考虑积分方程

$$u(x) = \lambda \int_a^x f(x, y, u(y)) \mathrm{d}y + v(x) \triangleq Fu(x), \qquad (4.2.16)$$

其中 f 如例 4.2.2(ii), $\lambda \in \mathbf{R}$ 与 $v \in C(J)$ 是给定的, $J = [a, b] (a < b)$. 方程(4.2.16)本身已具不动点方程的形式.取 $X = C(J)$, 则 $F: X \to X$. 利用例 4.2.2(ii)已得之结论并结合式(4.2.5)得

$$\mathrm{Lip} F \leqslant |\lambda| \beta(b-a).$$

因此,当 $|\lambda|$ 适当小时 $F: X \to X$ 是压缩映射.于是由定理 4.2.3 得出:当 $|\lambda|$ 适当小时方程(4.2.16)有惟一连续解.不过,以上论证实际上还是过于粗疏了.下面采用一个更细致的分析.任给 $u, v \in C(J), x \in J$, 有

① 将此处的方法与例 3.1.5 比较,你会发现二者非常类似;此处关于 $\mathrm{Lip} F^n$ 的不等式与例 3.1.5 所得的 $\| T^n \| \leqslant 1/n!$ 恰好相当.

$$|Fu(x) - Fv(x)| \leqslant |\lambda| \int_a^x |f(x,y,u(y)) - f(x,y,v(y))| \mathrm{d}y$$

$$\leqslant |\lambda| \int_a^x \beta \| u - v \|_0 \mathrm{d}y$$

$$= |\lambda| \beta(x - a) \| u - v \|_0.$$

作归纳假设

$$|F^{n-1}u(x) - F^{n-1}v(x)| \leqslant \frac{[|\lambda| \beta(x-a)]^{n-1}}{(n-1)!} \| u - v \|_0 \quad (n > 1),$$

则

$$|F^n u(x) - F^n v(x)| \leqslant |\lambda| \int_a^x \beta |F^{n-1}u(y) - F^{n-1}v(y)| \mathrm{d}y$$

$$\leqslant \frac{(|\lambda| \beta)^n}{(n-1)!} \int_a^x (y-a)^{n-1} \| u - v \|_0 \mathrm{d}y$$

$$= \frac{[|\lambda| \beta(x-a)]^n}{n!} \| u - v \|_0.$$

这就得出

$$\mathrm{Lip}F^n \leqslant [|\lambda| \beta(b-a)]^n / n! < 1 \quad (n \text{ 充分大}).$$

因此,限制 $|\lambda|$ 适当小是不必要的.

将方程(4.2.16)稍加变形,就可将其归入一般形式的方程:

$$\lambda x - Fx = y, \tag{4.2.17}$$

其中 $F: X \to X, \lambda \in \mathbf{R}$ 与 $y \in X$ 是给定的.如果 $F \in L(X)$,式(4.2.17)就成了上章中讨论过的方程(例如式(3.3.8)),其解法与谱理论有关.因此,通常将有关方程(4.2.17)的问题称为**非线性特征值问题**.若 $F = T \in L(X)$, $|\lambda| > \| T \|$,则方程(4.2.17)有惟一解 $x = (\lambda I - T)^{-1} y$ (参考 §3.1).下面是这一结果的非线性推广.

定理 4.2.6 若 $|\lambda| > \mathrm{Lip}F$,则逆算子 $(\lambda I - F)^{-1}: X \to X$ 存在,且

$$\mathrm{Lip}(\lambda I - F)^{-1} \leqslant (|\lambda| - \mathrm{Lip}F)^{-1}. \tag{4.2.18}$$

因此,方程(4.2.17)有惟一解 $x = (\lambda I - F)^{-1} y$,且 x 连续地依赖于 y.

证 仍然用前面已熟悉的方法:首先改写方程(4.2.17)为不动点方程

$$x = \lambda^{-1}(Fx + y) \triangleq Gx,$$

因 $\mathrm{Lip}G = |\lambda|^{-1}\mathrm{Lip}F < 1$，故 G 有惟一不动点 $x \in X$，x 即方程(4.2.17)的惟一解. 这表明 $\lambda I - F : X \to X$ 是双射，因此 $(\lambda I - F)^{-1} : X \to X$ 有定义.

余下证明式(4.2.18). 设 $y_i = (\lambda I - F)x_i, i = 1,2$，则

$$|\lambda| \| x_1 - x_2 \| = \| y_1 - y_2 + Fx_1 - Fx_2 \| \leqslant \| y_1 - y_2 \| + (\mathrm{Lip}F) \| x_1 - x_2 \|,$$

由此解出

$$\| x_1 - x_2 \| \leqslant \frac{\| y_1 - y_2 \|}{|\lambda| - \mathrm{Lip}F},$$

此不等式正好与式(4.2.18)相当. □

有趣的是，若 $F = T \in L(X)$，则不等式(4.2.18)成为

$$\|(\lambda I - T)^{-1}\| \leqslant (|\lambda| - \|T\|)^{-1},$$

此不等式可由式(3.1.10)直接推出. 以上事实及本节的其他一些结论表明，将线性算子理论中某些思想与结论移植于非线性算子，是可行的，或者至少是具有启发性的. 体会到这一点，对于非线性算子的学习不无好处.

§4.2.3 用于数值计算

我们已经看到，若 $\mathrm{Lip}F \leqslant r < 1$，则对任给 $x_0 \in D$，依式(4.2.9)构成的迭代序列 $\{x_n\}$ 都收敛于 F 在 D 中的惟一不动点 x^*. 因此，只要 n 适当大，就可用 x_n 作为 x^* 的近似值；由此而引起的误差可用式(4.2.10)或式(4.2.11)来估计. 这就为求方程的近似解提供了一个普遍方法，通常称为**迭代法**或**逐次逼近法**，它在数值计算中被广泛应用.

在将压缩映射原理应用于存在性证明时，我们只关心迭代序列的收敛性，而不关心其收敛速度. 在定理 4.2.4～4.2.6 中，甚至未直接提到逼近方程解的迭代序列. 至于数值计算，则完全不同：我们不仅关心迭代序列的构成，更关心其收敛速度. 误差估计式(4.2.10)表明，迭代序列 $\{x_n\}$ 的收敛速度不慢于几何级数 $\sum r^n$ 的收敛速度. 因此，构成迭代序列的一个首要要求是 r 应尽可能地小. 这就应这样构成与原方程等价的不动点方程 $x = Fx$，使得 $\mathrm{Lip}F$ 尽可能小；而且随着 $\rho \to 0$，$\mathrm{Lip}(F \mid B_\rho(x^*))$ 将越来越小(从而迭代序列收敛得越来越快)，最好是

$$\lim_{\rho \to 0} \mathrm{Lip}(F \mid B_\rho(x^*)) = 0. \tag{4.2.19}$$

若 F 是 C^1 映射，则由不等式(4.2.6)推出，当 $F'(x^*) = 0$ 时条件(4.2.19)式必满足.

下面通过两个具体例子来说明.

例 4.2.7(Jacobi 迭代法) 设 $A=[a_{ij}]\in\mathbf{R}^{n\times n}$, $b\in\mathbf{R}^n$. 为用迭代法求解线性方程

$$Ax=b, \tag{4.2.20}$$

分解 $A=D+Q$, 其中 $D=\operatorname{diag}(a_{11},a_{22},\cdots,a_{nn})$, Q 是将 A 的对角线中的元变为零而得之矩阵, 即 $Q=A-D$, 则方程(20)可依如下方式转化为不动点方程:

$$Dx=-Qx+b,$$

$$x=-D^{-1}Qx+D^{-1}b\triangleq Bx+g. \tag{4.2.21}$$

要使方程(4.2.21)适于用来构造迭代序列, 要求 $\|B\|<1$ (这并非必要条件; 由定理 3.1.4, 只要 $r_\sigma(B)<1$ 就能使迭代序列收敛), 最好是 $\|B\|$ 尽可能地小. 注意到(依命题 2.2.1)

$$\|B\|_\infty=\max_i\sum_{j\neq i}|a_{ij}|/|a_{ii}|$$

$$=\max_i|a_{ii}|^{-1}\sum_{j\neq i}|a_{ij}|,$$

因此, 只要

$$\sum_{j\neq i}|a_{ij}|<|a_{ii}|\quad(1\leqslant i\leqslant n), \tag{4.2.22}$$

就有 $\|B\|_\infty<1$, 从而保证迭代序列收敛. 满足条件(4.2.22)式的矩阵 A 称为**按行严格对角占优矩阵**. 取 $r=\|B\|_\infty$, 就可应用误差估计公式(4.2.10)或式(4.2.11). 不过应注意, 与对 B 取范数 $\|B\|_\infty$ 相对应, 应用式(4.2.10)或式(4.2.11)时, 对向量 x 要用范数 $\|x\|_\infty$.

类似的方法, 亦可用于由无穷矩阵表示的无穷线性方程组(参考命题 2.2.2), 且可得类似的结论.

例 4.2.8(Newton 迭代法) 设 $F:\Omega\subset X\to X$ 是 C^2 映射, 今要解方程

$$Fx=0. \tag{4.2.23}$$

若预先估计到方程(4.2.23)有解 x^*, 且 $F'(x^*)$ 可逆, 则在 x^* 邻近 $F'(x)$ 亦可逆, 因此可将方程(4.2.23)转化为不动点方程

$$x=x-[F'(x)]^{-1}Fx\triangleq Gx. \tag{4.2.24}$$

一个稍细致的分析可以得出

$$G'(x)h=F'(x)^{-1}(F''(x)h)F'(x)^{-1}Fx,$$

这表明 $G'(x^*)=0$，从而在 x^* 邻近 Lip$G$ 必充分小. 因此，如果依式(4.2.24)来构成迭代序列，则可能达到较高的收敛速度. 惟一成问题的是：为保证迭代序列收敛，应选取 x_0 适当接近于待求的 x^*；然而 x^* 却是未知的. 因此在选择 x_0 时要用到关于映射 F 的适当条件，这些可在有关数值计算的专业书中找到.

看一个最简单的例子. $\sqrt[3]{2}$是方程

$$F(x)=x^3-2=0$$

的解. 令

$$G(x)=x-\frac{F(x)}{F'(x)}=\frac{2}{3}\left(x+\frac{1}{x^2}\right).$$

取 $x_0=1$，则经 4 次迭代后得出：

$$x_1=4/3,\quad x_2=91/72,$$

$$x_3=1.25993349,$$

$$x_4=1.25992105,$$

其中 x_4 的 9 位数字全为准确数字.

§4.3 隐函数定理

在上节中我们看到，压缩映射原理可用于解决各种类型的方程问题. 隐函数方程就是应用上最常见的方程之一. 本节应用压缩映射原理来建立 Banach 空间中的隐函数定理，它无疑是无限维分析学的最重要的成果之一.

§4.3.1 反函数定理

反函数定理与隐函数定理可互相推出. 但从形式上看，反函数定理显得简单一些，因此我们首先建立反函数定理，然后由其推出隐函数定理.

若 $f:(a,b)\subset\mathbf{R}\to\mathbf{R}$ 是连续可微函数，$f'(x_0)\neq 0$，$x_0\in(a,b)$，则 f 在某个含 x_0 的区间 J 上是严格单调函数，因而 f 作为 J 上的函数存在反函数 $f^{-1}(y)$，$f^{-1}(y)$ 也是连续可微函数，且

$$f^{-1\prime}(y)=1/f'(x)\quad(y=f(x)).$$

类比于此，对于 C^1 映射 $F:\Omega\subset X\to Y$，若 $x_0\in\Omega$，$F'(x_0)$ 可逆，能否断定 F 在 x_0 邻近有 C^1 的逆映射 F^{-1}，且

$$F^{-1'}(y) = F'(x)^{-1} \quad (y = Fx).$$

反函数定理正是对上述问题给出一个肯定回答.

定理 4.3.1(反函数定理)　设 X, Y 是 Banach 空间, $F: \Omega \subset X \to Y$ 是一 $C^r(1 \leqslant r \leqslant \infty)$ 映射, $x_0 \in \Omega, F'(x_0): X \cong Y$ ($\cong$表同构,下同). 则存在 x_0 的开邻域 U 与 Fx_0 的开邻域 V, 使得 $F: U \to V$ 是双射,其逆映射 $F^{-1}: V \to U$ 是 C^r 映射,且

$$(F^{-1})'(Fx) = (F'(x))^{-1} \quad (x \in U). \tag{4.3.1}$$

证　证明稍长,分为几个步骤进行.

(i) 简化问题. 令 $T = F'(x_0)$, 由逆算子定理知 T^{-1} 有界. 定义

$$Gx = T^{-1}(F(x_0 + x) - Fx_0),$$

则 $G: \Omega - x_0 \subset X \to X$ 是一 C^r 映射, $G(0) = 0, G'(0) = I$. 若在 $x = 0$ 邻近 G 有 C^r 的逆映射 G^{-1}, 则不难验证

$$F^{-1}y = x_0 + G^{-1}T^{-1}(y - Fx_0)$$

就是 F 在 x_0 邻近的 C^r 的逆映射. 因此,不失一般性,不妨一开始就设 $X = Y$, $x_0 = 0 \in \Omega, F(0) = 0, F'(0) = I$.

(ii) 证逆映射的存在性,即确定 $x = 0$ 的开邻域 U 与 V, 使得 $\forall y \in V$, 方程 $Fx = y$ 有惟一解 $x \in U$. 因这是一个方程问题,故如上节中所作的,将 $Fx = y$ 转化为不动点方程

$$x = x - Fx + y \triangleq Gx.$$

我们试图在某个球 $\bar{B}_\delta(0)$ 上对 G 应用压缩映射原理, δ 待定. $\forall x, z \in \bar{B}_\delta(0)$, 有

$$\begin{aligned} \|Gx\| &\leqslant \|x - Fx\| + \|y\| \\ &\leqslant \sup_{0 \leqslant t \leqslant 1} \|I - F'(tx)\| \|x\| + \|y\|; \text{(用式(4.1.10))} \end{aligned} \tag{4.3.2}$$

$$\begin{aligned} \|Gx - Gz\| &= \|Fx - Fz - (x - z)\| \\ &\leqslant \sup_{s \in [x,z]} \|F'(s) - I\| \|x - z\|. \text{(用式(4.1.10))} \end{aligned} \tag{4.3.3}$$

如果我们希望 $G\bar{B}_\delta(0) \subset \bar{B}_\delta(0), \mathrm{Lip}(G \mid \bar{B}_\delta(0)) \leqslant 1/2$, 就应这样选择 δ 与 y: 取 $\delta > 0$, 使当 $\|x\| \leqslant \delta$ 时 $\|F'(x) - I\| < 1/2$, 且 $F'(x) \in GL(X)$, 而 $y \in V \triangleq B_{\delta/2}(0)$. 因 $F'(0) = I$ 而 $F'(x)$ 连续,故如上的 δ 必存在. 于是对 $x, z \in \bar{B}_\delta(0)$, 由式(4.3.2)推出 $\|Gx\| < \delta$; 由式(4.3.3)推出

$$\|Gx - Gz\| \leqslant (1/2)\|x - z\|.$$

于是由压缩映射原理得出：G 在 $\bar{B}_\delta(0)$ 上有惟一不动点 x，因而 $Fx=y$. 因 $\|x\|=\|Gx\|<\delta$，故 $x\in U\triangleq B_\delta(0)\cap F^{-1}V$. 因此 $F:U\to V$ 为双射，U 与 V 都是 0 的开邻域.

(iii) 证 $F^{-1}:V\to U$ 连续. 设 $y_i=Fx_i\in V, x_i\in U, i=1,2$，则由式(4.3.3)有

$$\frac{1}{2}\|x_1-x_2\|\geqslant\|x_1-x_2-(Fx_1-Fx_2)\|$$
$$\geqslant\|x_1-x_2\|-\|y_1-y_2\|,$$

故

$$\|F^{-1}y_1-F^{-1}y_2\|\leqslant 2\|y_1-y_2\|\quad(y_1,y_2\in V).\qquad(4.3.4)$$

(iv) 证式(4.3.1). 取定 $x\in U$，设 $y=Fx$、$x+h\in U, y+k=F(x+h)\in V$. 由式(4.3.4)有 $\|h\|\leqslant 2\|k\|$；

$$\|F^{-1}(y+k)-F^{-1}y-(F'(x))^{-1}k\|$$
$$=\|(F'(x))^{-1}[F'(x)h-k]\|$$
$$\leqslant\|(F'(x))^{-1}\|\,\|F(x+h)-Fx-F'(x)h\|$$
$$=o(\|h\|)=o(\|k\|),\quad(\text{用 }\|h\|\leqslant 2\|k\|)$$

这正表明式(4.3.1)成立.

(v) 证 $F^{-1}\in C^r$. 显然只需要考虑 $r<\infty$ 的情况. 从已证的式(4.3.1)直接看出 $(F^{-1})'(y)$ 对 y 连续，因此 $F^{-1}\in C^1$. 将 $(F^{-1})'(y)$ 作如下分解：

$$y\xrightarrow{F^{-1}}x\xrightarrow{F'}F'(x)\to(F'(x))^{-1}=(F^{-1})'(y),$$

于是利用 $F'\in C^{r-1}$ 与归纳假设 $F^{-1}\in C^{r-1}$ 得出 $(F^{-1})'\in C^{r-1}$，从而 $F^{-1}\in C^r$ 得证. □

仔细检查上述证明你会发现，我们隐蔽地利用了以下结论：

(A) 若 $T\in L(X)$，$\|T-I\|<1/2$，则 T 可逆.

(B) C^r 映射的复合映射是 C^r 映射.

(C) $\mathrm{GL}(X)\to L(X), T\to T^{-1}$ 是 C^∞ 映射.

结论(A)可从引理 3.1.1 推出(引理 3.1.1 亦适用于实 Banach 空间). 至于结论(B)与(C)，在有限维空间中似乎不成问题，但在无限维空间中并不显然，因而需要严格证明. 如果你对于这一类的结论并无进一步的兴趣，可以不看下面补证的两条引理.

引理 4.3.2　设 $F:\Omega\subset X\to Y$ 与 $G:\Omega'\subset Y\to Z$ 是 C^r 映射，$0\leqslant r\leqslant\infty$，$F\Omega\subset\Omega'$，则 $H=G\circ F\in C^r$. 简言之，C^r 映射的复合映射是 C^r 映射.

证　显然只需考虑 $r<\infty$ 的情况. 当 $r=0$ 时引理结论显然成立. 下面设 $r\geqslant 1$，且设引理已对 C^{r-1} 映射证明. 映射 $H'(x)=G'(Fx)F'(x)$ 可作如下分解：

$$x\to(G'(Fx),F'(x))\to G'(Fx)F'(x). \tag{4.3.5}$$

由归纳假设，$x\to G'(Fx)$ 是 C^{r-1} 映射，因而(用定理 4.1.7)

$$x\to(G'(Fx),F'(x))$$

是 C^{r-1} 映射. 其次，直接看出

$$L(Y,Z)\times L(X,Y),\quad (A,B)\to AB$$

为 C^∞ 映射(实际上，此映射的 3 阶导数恒为零). 于是将归纳假设用于式(4.3.5)得 $H'\in C^{r-1}$，从而 $H\in C^r$. □

引理 4.3.3　令 $G=\mathrm{GL}(X)$（X 上的拓扑自同构之全体)，则 G 是 $L(X)$ 的开子集，且**反演映射**

$$J:G\to G,\quad A\to A^{-1} \tag{4.3.6}$$

是 C^∞ 映射.

证　采用 §3.1.1 中的方法. 取定 $A\in G$. 若 $B\in L(X)$，$\|B\|$ 充分小，则有

$$\begin{aligned}(A+B)^{-1}&=A^{-1}(I+BA^{-1})^{-1}\\&=\sum_{n=0}^{\infty}A^{-1}(-BA^{-1})^n,\quad(\text{参照式}(3.1.6))\end{aligned}$$

这表明 $A+B\in G$，且

$$\begin{aligned}&\|(A+B)^{-1}-A^{-1}+A^{-1}BA^{-1}\|\\&=\Big\|\sum_{n=2}^{\infty}A^{-1}(-BA^{-1})^n\Big\|\\&\leqslant\sum_{n=2}^{\infty}\|A^{-1}\|^{n+1}\|B\|^n=o(\|B\|).\end{aligned}$$

以上估计表明

$$J'(A)B=-A^{-1}BA^{-1}\triangleq\Phi(B), \tag{4.3.7}$$

$$\Phi:L(X)\to L(X),\quad B\to A^{-1}BA^{-1}$$

显然为有界线性算子.令

$$L_A : L(X) \to L(X), \quad B \to AB,$$

$$R_A : L(X) \to L(X), \quad B \to BA,$$

则 L_A 与 R_A 均为有界线性算子,而 $A \to L_A$ 与 $A \to R_A$ 则是从 $L(X)$ 到 $L(L(X))$ 的有界线性算子.由式(4.3.7)有

$$J'(A) = -L_{J(A)} \circ R_{J(A)}. \tag{4.3.8}$$

作归纳假设: $J \in C^{r-1}(r \geqslant 1)$,则由式(4.3.8)及引理 4.3.2 得 $J' \in C^{r-1}$,从而 $J \in C^r$.因 r 是任意的,故 $J \in C^\infty$. □

注意引理 4.3.3 正是以下初等结论的推广: $\varphi(x) = 1/x(0 \neq x \in \mathbf{R})$ 无限次可微,且 $\varphi'(x) = -\varphi^2(x)$.

§4.3.2 隐函数定理

首先让我们回忆一下微积分学中的隐函数定理.设 $F : \Omega \subset \mathbf{R}^2 \to \mathbf{R}$ 是一个连续可微函数,$(x_0, y_0) \in \Omega$,$F(x_0, y_0) = 0$,$F_y(x_0, y_0) \neq 0$,则在点 (x_0, y_0) 邻近可从方程 $F(x, y) = 0$ 解出 $y = y(x)$,使得 $y_0 = y(x_0)$,

$$y'(x) = -F_x(x, y(x))/F_y(x, y(x)).$$

可庆幸的是,以上结论可推广于 Banach 空间中的隐函数方程.实际上,推广后的隐函数定理只不过是定理 4.3.1 的一个推论.

定理 4.3.4(隐函数定理) 设 X, Y, Z 是 Banach 空间,$F : \Omega \subset X \times Y \to Z$ 是一 C^r 映射 $(1 \leqslant r \leqslant \infty)$,$(x_0, y_0) \in \Omega$,$F(x_0, y_0) = 0$,$F_y(x_0, y_0) : Y \cong Z$.则存在点 (x_0, y_0) 的开邻域 $U \times V (\subset \Omega)$ 及惟一的 C^r 映射 $f : U \to V$,使得

$$F(x, f(x)) = 0 (x \in U), \quad f(x_0) = y_0; \tag{4.3.9}$$

$$f'(x) = -[F_y(x, f(x))]^{-1} F_x(x, f(x)) \quad (x \in U). \tag{4.3.10}$$

证 作如下辅助映射:

$$G : \Omega \to X \times Z, \quad (x, y) \to (x, F(x, y)).$$

由定理 4.1.7,有 $G \in C^r$.约定 $F_x^0 = F_x(x_0, y_0)$,F_y^0 仿此,则

$$G'(x_0, y_0) = \begin{bmatrix} I & 0 \\ F_x^0 & F_y^0 \end{bmatrix}.$$

仿照线性代数中求逆矩阵的方法,形式地求出

$$[G'(x_0,y_0)]^{-1} = \begin{bmatrix} I & 0 \\ -(F_y^0)^{-1}F_x^0 & (F_y^0)^{-1} \end{bmatrix} \triangleq Q; \tag{4.3.11}$$

然后由直接计算验证 JQ 与 QJ 均为单位算子,此处 $J = G'(x_0,y_0)$. 因此式(4.3.11)事实上为真,从而 $G'(x_0,y_0): X\times Y \cong X\times Z$. 将反函数定理用于 G 得出:G 在 (x_0,y_0) 的某个开邻域 $U\times V$ 上存在 C^r 的逆映射 G^{-1}, G^{-1} 定义在 $(x_0,0)\in X\times Z$ 的某个开邻域内,且 $G^{-1}(x_0,0)=(x_0,y_0)$. 设 $P: X\times Y\to Y$ 是投影,令

$$f(x) = PG^{-1}(x,0),$$

则 $f: U\to V$ 是 C^r 映射, $f(x_0)=y_0$. 由

$$(x,F(x,f(x))) = G(x,f(x)) = (x,0)$$

推出 $F(x,f(x))=0(x\in U)$,可见 $f(x)$ 满足式(4.3.9). 微分恒等式 $F(x,f(x))=0$ (应用式(4.1.17)')得

$$F_x(x,f(x))h + F_y(x,f(x))f'(x)h = 0 \quad (h\in X),$$

由此得出(10). 由逆映射的惟一性推出如上 f 的惟一性. □

看看定理 4.3.4 的有限维形式是有益的. 在定理 4.3.4 中取 $X=\mathbf{R}^n$, $Y=Z=\mathbf{R}^m$ (不必 $m\neq n$)得到:若 $\Omega\subset\mathbf{R}^{n+m}$, $F_i(1\leqslant i\leqslant m)$ 是 Ω 内的实值 C^r 函数, $(x^0,y^0)\in\Omega$, $F_i(x^0,y^0)=0(1\leqslant i\leqslant m)$,

$$\det[\partial F_i(x^0,y^0)/\partial y_k]\neq 0,$$

则在点 (x^0,y^0) 邻近可从方程组

$$F_i(x_1,\cdots,x_n,y_1,\cdots,y_m)=0 \quad (1\leqslant i\leqslant m) \tag{4.3.12}$$

惟一地解出

$$y_i = f_i(x_1,x_2,\cdots,x_n) \quad (1\leqslant i\leqslant m),$$

f_i 是点 $x^0=(x_1^0,x_2^0,\cdots,x_n^0)^{\mathrm{T}}$ 邻近的 C^r 函数,它满足

$$y_i^0 = f_i(x_1^0,x_2^0,\cdots,x_n^0), \quad y^0=(y_1^0,y_2^0,\cdots,y_n^0)^{\mathrm{T}};$$

$$\left[\frac{\partial f_i(x)}{\partial x_j}\right] = -\left[\frac{\partial F_i(x,y)}{\partial y_k}\right]^{-1}\left[\frac{\partial F_k(x,y)}{\partial x_j}\right],$$

其中 $x=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$, $y=(y_1,y_2,\cdots,y_m)^{\mathrm{T}}$.

你不妨设想一下,如果我们不用 Banach 空间的语言,而保持平常的坐标形式,

关于方程组(4.3.12)的可解性证明,是否容易理出一个头绪来.这个例子大概让你信服地证明了:泛函分析的抽象方法常常能使一些复杂的问题明显简化.

§4.3.3 某些应用

隐函数定理(包括反函数定理)在多个数学领域有大量深刻的应用.许多应用涉及较深入的专门知识,此处仅能举最简单的例子,它们亦颇能说明问题.

首先作一般说明.设某一问题同时涉及量 x,y. 此处我们在尽可能一般的意义上理解量,它可以是很复杂的对象(如函数或算子),而并不限于指数量或有限维向量.如果能根据问题的条件得出 x 与 y 之间的某个约束关系:

$$F(x,y)=0,$$

那么我们希望由此知道:

(A) 是否存在函数关系 $y=f(x)$?

(B) 能否从 F 的一定可微性推出 $y=f(x)$ 的同样可微性?

而隐函数定理恰好可用来回答以上问题.在很多问题中,函数关系 $y=f(x)$ 存在及其可微性是重要的,而 $f(x)$ 的具体表达式则未必重要.而且应指出,隐函数定理对于具体解出 $y=f(x)$ 并无帮助.有时,函数关系 $y=f(x)$ 是已知的,应用隐函数定理的惟一作用是得出 $f(x)$ 的某种可微性,即解决上述的问题(B).即使仅仅在这个意义上应用隐函数定理,其意义亦不可低估,因很少有其他分析定理能起类似的作用.

现在来看几个具体例子.

例 4.3.5(特征值问题) 设 λ_0 是 $A_0\in\mathbf{C}^{n\times n}$ 的单重特征值.则当 $A\in\mathbf{C}^{n\times n}$ 充分接近于 A_0 时, A 亦有单重特征值 λ_A, λ_A 充分接近于 λ_0, 且 λ_A 关于 A 为 C^∞ 函数.

证 首先给出 $A\in\mathbf{C}^{n\times n}$ 与其特征值 λ 之间的约束关系,这就是

$$F(A,\lambda)\triangleq\det(\lambda I-A)=0. \tag{4.3.13}$$

显然, $F:\mathbf{C}^{n\times n}\times\mathbf{C}\to\mathbf{C}$ 是 C^∞ 函数, $F(A_0,\lambda_0)=0$. 因 λ_0 是 n 次多项式 $P(\lambda)\triangleq F(A_0,\lambda)$ 的单重零点,故必

$$F_\lambda(A_0,\lambda_0)=P'(\lambda_0)\neq0.$$

于是由隐函数定理得出:在 (A_0,λ_0) 邻近可从方程(4.3.13)解出 C^∞ 函数 $\lambda=\lambda_A$, λ_A 是 A 的特征值.因当 $A\approx A_0$ 时

$$F_\lambda(A,\lambda_A)\approx F_\lambda(A_0,\lambda_0)\neq0,$$

故 λ_A 是 A 的单重特征值. □

你试设想一下,如果不用隐函数定理,该如何解决例 4.3.5 中的问题.须知,我们无法求得用 A 表示 λ_A 的明显表达式.你必定觉得,A 作为函数 λ_A 的"自变量",似乎太复杂了.而对于应用隐函数定理,这是完全不必介意的事.

一个类似于例 4.3.5 但更具一般性的应用是:

例 4.3.6(方程的扰动)　考虑 $\mathbf{R}^n$ 中的非线性方程

$$u(x) = 0. \tag{4.3.14$_u$}$$

函数 u 通常由实际资料确定,未必完全准确.那么 u 的误差对于方程 $(4.3.14)_u$ 的解影响如何? 这无疑是不可忽视的问题.为应用隐函数定理,将方程$(4.3.14)_u$ 改写成

$$F(u,x) \triangleq u(x) = 0,$$

其中 $u \in C^1(\bar{\Omega},\mathbf{R}^n)$, $x \in \Omega$, $\Omega \subset \mathbf{R}^n$ 是某个有界开区域,在 $C^1(\bar{\Omega},\mathbf{R}^n)$ 中使用范数

$$\| u \| = \sup_{x\in\bar{\Omega}} |u(x)| + |u'(x)|.$$

容易看出,$F: C^1(\bar{\Omega},\mathbf{R}^n)\times\Omega \to \mathbf{R}^n$ 是 C^1 映射(注意 $F(u,x)$ 对 u 是有界线性算子).给定 $u_0 \in C^1(\bar{\Omega},\mathbf{R}^n)$,若方程 $(4.3.14)_{u_0}$ 有解 $x_0 \in \Omega$ 且 $u'_0(x_0) \in GL(\mathbf{R}^n)$,则

$$F_x(u_0,x_0) = u'_0(x_0): \mathbf{R}^n \cong \mathbf{R}^n.$$

于是由隐函数定理推出:对于 u_0 的 C^1 近似 u,方程 $(4.3.14)_u$ 在 x_0 邻近有惟一解 x_u,且 x_u 是 u 的 C^1 函数;特别,x_u 是 u 的连续函数,即 u 的微小变动仅导致方程 $(4.3.14)_u$ 解的微小变动.这一结论无疑是令人满意的.

例 4.3.7(到达时间)　设 H 是一实 Hilbert 空间,$f: X\to\mathbf{R}$ 是 C^1 函数,$J=[a,b](a<b)$,当 $x\in f^{-1}(J)$ 时 $\nabla f(x)\neq 0$. C^1 映射 $\sigma: \mathbf{R}\times X\to X$ 满足:

$$\sigma(0,x) = x, \quad \frac{\mathrm{d}}{\mathrm{d}t}\sigma(t,x) = -\nabla f(\sigma(t,x)).$$

这意味着 $\sigma(t,x)$ 是向量场 $-\nabla f(x)$ 的**流**.设 $\forall x\in f^{-1}(J)$,有惟一 $t_x\in\mathbf{R}$,使得 $f(\sigma(t_x,x)) = a$ (t_x 称为沿 $\nabla f(x)$ 的轨线到达水平集 $\{f=a\}$ 的时间,简称为**到达时间**).这意味着,$t=t_x$ 是方程

$$F(x,t) \triangleq f(\sigma(t,x)) - a = 0$$

的惟一解.因 $F: X\times\mathbf{R}\to\mathbf{R}$ 为 C^1 函数,且当 $f(\sigma(t,x)) = a$ 时

$$F_t(x,t) = \langle \nabla f(\sigma(t,x)), \frac{\mathrm{d}}{\mathrm{d}t}\sigma(t,x)\rangle$$

$$= - \| \nabla f(\sigma(t,x)) \|^2 < 0,$$

故由隐函数定理得出：t_x 是 x 的 C^1 函数.特别，t_x 连续地依赖于 x；若 x 有微小变动，则到达时间亦只有微小变动.

你注意到，在例 4.3.7 中，到达时间 t_x 的存在性是预先假定的(存在性可由其他理由确立)，因而应用隐函数定理的惟一目的是获得函数 $x \to t_x$ 的可微性.实际上，只要能确定对某个 x_0，到达时间 t_{x_0} 存在，就可由隐函数定理推出，至少在 x_0 邻近，到达时间 t_x 是存在的.

例 4.3.8(孤立奇点) 设 $F: \Omega \subset X \to X$ 是一个 C^1 映射，通常称 F 为 Ω 内的**向量场**，称 F 的零点(即方程 $Fx = 0$ 的解)为向量场的**奇点**.奇点概念在动力系统理论中具有基本意义.现在考虑这样的问题：若 $x_0 \in \Omega$ 是 F 的奇点，$F'(x_0) \in GL(X)$ (这样的 x_0 姑且称为**正则奇点**)，F 在 x_0 的充分小的邻域内是否有其他奇点？回答是明确的：必存在 x_0 的邻域 U，x_0 是 F 在 U 内的惟一奇点.事实上，由反函数定理，有 x_0 的邻域 U，使 F 在 U 上是可逆的，因而方程 $Fx = 0$ 在 U 内只能有一个解，即 x_0.这样，我们就得出结论：C^1 向量场的正则奇点必为孤立奇点.

以上结论特别可用于 n 维自治微分系统

$$\frac{\mathrm{d}x}{\mathrm{d}t} = F(x), \tag{4.3.15}$$

其中 $F: \Omega \subset \mathbf{R}^n \to \mathbf{R}^n$ 是 C^1 映射.若 $x_0 \in \Omega$，$F(x_0) = 0$，$\det F'(x_0) \neq 0$ (你记得 $F'(x_0)$ 就是 F 在 x_0 的 Jacobi 矩阵)，则 x_0 是式(4.3.15)的孤立奇点.

§4.4 紧算子

在§3.3 中我们看到，紧线性算子具有一系列良好的性质，在许多方面颇接近于有限维空间上的线性算子.一个自然的问题是，能否在非线性算子中界定一类非线性紧算子，使之具有较好的性质，例如可与有限维空间上的连续映射相比拟？在本节中你将看到，对上述问题有一个令人满意的回答.

§4.4.1 紧算子及其性质

以下定义是定义 3.3.1 的一个直接推广.

定义 4.4.1 设 $F: D \subset X \to Y$ 是一连续映射.若对任一有界序列 $\{x_n\} \subset$

D, $\{Fx_n\}$ 恒有收敛子列,则称 F 为**紧算子或紧映射**[①].

除了预设 F 连续且不必 $D = X$ 之外,上述定义无非是定义 3.3.1 的翻版. 因此毫不足怪,紧算子将有某些类似于紧线性算子的性质. 首先可以指出,如同线性算子一样,连续映射 $F: D \subset X \to Y$ 为紧算子的充要条件是: F 将每个有界集 $B \subset D$ 映成 Y 中的相对紧集;若 D 本身有界,则 $F \in C(D, Y)$ 是紧算子 $\Leftrightarrow FD$ 是相对紧集. 由此容易说明,对于 $F \in C(D, Y)$, 当 $\dim Y < \infty$ 时, F 是紧算子 $\Leftrightarrow F$ 映有界集为有界集;当 $\dim X < \infty$ 且 D 为闭集时, $F \in C(D, Y)$ 必为紧算子. 可见,只有在无限维空间中,紧算子概念才有实质意义.

例 3.3.4 指出了某些常用的线性积分算子是紧算子. 现在指出,类似地构成的非线性积分算子亦是紧算子. 这一事实对于紧算子的应用意义颇大.

例 4.4.2 设 $J = [a, b] (a < b)$, $f \in C(J \times J \times \mathbf{R}, \mathbf{R})$. 利用函数 f 可定义两个形式上相近的积分算子,分述如下.

(i) 首先考虑积分算子

$$Fu(x) = \int_a^b f(x, y, u(y)) \mathrm{d}y \quad (u \in C(J)). \tag{4.4.1}$$

由 f 的连续性直接推出, $\forall u \in C(J)$, 有 $Fu \in C(J)$. 今指明 $F: C(J) \to C(J)$ 为紧算子. 首先说明 F 连续. 设 $\{u_n\} \subset C(J)$, $u_n \rightrightarrows u (n \to \infty)$, 今要证 $Fu_n \rightrightarrows Fu$. 令 $\beta = \sup\limits_n \| u_n \|_0$. 因 f 在 $J \times J \times [-\beta, \beta]$ 上一致连续, $\forall \varepsilon > 0$, 可取 $\delta > 0$, 使对 $x, x', y, y' \in J$, $z, z' \in [-\beta, \beta]$ 有

$$\begin{aligned} &\max\{|x - x'|, |y - y'|, |z - z'|\} < \delta \\ &\Rightarrow |f(x, y, z) - f(x', y', z')| < \varepsilon. \end{aligned} \tag{4.4.2}$$

取 $N > 0$, 使当 $n \geqslant N$ 时 $\| u_n - u \|_0 < \delta$. 于是当 $n \geqslant N$, $x \in J$ 时有

$$\begin{aligned} |Fu_n(x) - Fu(x)| &\leqslant \int_a^b |f(x, y, u_n(y)) - f(x, y, u(y))| \mathrm{d}y \\ &\leqslant \varepsilon(b - a). \end{aligned}$$

可见 $Fu_n \rightrightarrows Fu (n \to \infty)$, 因此 F 连续.

其次证 F 满足紧性条件. 任给有界集 $B \subset C(J)$, 今要证 FB 有界且等度连续,从而相对紧(参看定理 1.4.6). 令 $\beta = \sup\limits_{u \in B} \| u \|_0$,

① 从与定义 3.3.1 对照来看,这一术语是很自然的. 不过,文献中用语颇不一致. 一些作者将这里所说的紧算子称为**全连续算子**,仅当 $\overline{FD}$ 紧时称 F 为紧算子.

$$M = \sup\{|f(x,y,z)|: x,y \in J, |z| \leqslant \beta\}. \qquad (4.4.3)$$

$\forall u \in B$, 有

$$\|Fu\|_0 \leqslant \sup_{x \in J}\int_a^b |f(x,y,u(y))| \mathrm{d}y \leqslant M(b-a),$$

可见 FB 有界. 设 ε, δ 如式(4.4.2), 则当 $x, x' \in J$, $|x - x'| < \delta$ 时, 对任何 $u \in B$ 有

$$|Fu(x) - Fu(x')| \leqslant \int_a^b |f(x,y,u(y)) - f(x',y,u(y))| \mathrm{d}y$$

$$\leqslant \varepsilon(b-a),$$

可见 FB 等度连续. 这就验证了 F 为紧算子.

(ii) 现在考虑稍不同于式(4.4.1)的积分算子

$$Fu(x) = \int_a^x f(x,y,u(y))\mathrm{d}y \quad (x \in J, u \in C(J)). \qquad (4.4.4)$$

同样可以验证 $F: C(J) \to C(J)$ 为紧算子, 证明大体上类似于(i), 只是对 FB($B \subset C(J)$ 有界)等度连续的验证略有差别: 设 $\beta, M, \varepsilon, \delta$ 如式(4.4.2), 式(4.4.3)且 $\delta < \varepsilon$, 则当 $x, x' \in J$, $0 \leqslant x' - x < \delta$ 时, 对任给 $u \in B$ 有

$$|Fu(x) - Fu(x')| \leqslant \int_a^x |f(x,y,u(y)) - f(x',y,u(y))| \mathrm{d}y$$

$$+ \int_x^{x'} |f(x',y,u(y))| \mathrm{d}y$$

$$\leqslant \varepsilon(b-a) + \delta M < \varepsilon(b-a+M),$$

这表明 FB 等度连续.

以下结果建立了紧算子与紧线性算子之间的联系.

命题 4.4.3 设 $F: D \subset X \to Y$ 是一个紧算子, x 是 D 的内点, F 在 x 可微, 则 $F'(x) \in CL(X,Y)$.

证 令 $T = F'(x)$, 只需证 $TB_1(0)$ 全有界(用定理 1.4.3, 注意本章假定 Y 完备). $\forall \varepsilon > 0$, 取 $\delta > 0$, 使当 $h \in X$, $\|h\| < \delta$ 时

$$\|F(x+h) - Fx - Th\| \leqslant \varepsilon \|h\| < \varepsilon\delta,$$

因而

$$TB_\delta(0) \subset FB_\delta(x) - Fx + B_{\varepsilon\delta}(0). \qquad (4.4.5)$$

因 $FB_\delta(x)$ 相对紧, 故有有限集 $\{y_i\} \subset Y$, 使得

$$FB_\delta(x) \subset \bigcup_i B_{\varepsilon\delta}(y_i). \tag{4.4.6}$$

于是

$$\begin{aligned} TB_1(0) &= \delta^{-1} TB_\delta(0) \subset \delta^{-1}[FB_\delta(x) - Fx + B_{\varepsilon\delta}(0)] \quad (\text{用式(4.4.5)}) \\ &\subset \delta^{-1}[\bigcup_i B_{\varepsilon\delta}(y_i) + B_{\varepsilon\delta}(0) - Fx] \quad (\text{用式(4.4.6)}) \\ &\subset \bigcup_i B_{2\varepsilon}(z_i) \quad (z_i = \delta^{-1}(y_i - Fx)), \end{aligned}$$

可见 $TB_1(0)$ 全有界,如所要证. □

本节引言中曾预期紧算子接近于有限维空间中的连续映射,以下定理为证实这一点走出了第一步.

定理 4.4.4(逼近定理)　设 $F: D \subset X \to Y$ 是紧算子, D 有界. 则 $\forall \varepsilon > 0$, 存在 $G \in C(D, Y)$, 使得 $\| Gx - Fx \| \leqslant \varepsilon (\forall x \in D)$, 且 GD 含于 Y 的某个有限维子空间.

证　取定 $\varepsilon > 0$. 因 FD 全有界,故有有限集 $\{y_i\} \subset Y$, 使得

$$FD \subset \bigcup_i B_\varepsilon(y_i).$$

不妨设 $\{y_i\} \subset FD$ (何故?). 下面将 Gx 定义为 y_i 的一个凸组合:

$$Gx = \sum_i q_i(x) y_i \quad (x \in D),$$

其中

$$q_i(x) = p_i(x) / \sum_j p_j(x), \quad p_i(x) = d(Fx, Y \setminus B_\varepsilon(y_i)).$$

显然 $p_i(\cdot)$ 在 D 上连续(参考第一章习题 1), 当 $Fx \in B_\varepsilon(y_i)$ 时 $p_i(x) > 0$. 由 $FD \subset \bigcup_i B_\varepsilon(y_i)$ 推出 $\sum p_i(x) > 0 (\forall x \in D)$, 因此 $q_i(\cdot)$ 亦在 D 上连续, 且 $\sum q_i(x) = 1 (\forall x \in D)$. 注意到

$$q_i(x) \neq 0 \Rightarrow p_i(x) \neq 0 \Rightarrow Fx \in B_\varepsilon(y_i),$$

故对任给 $x \in D$ 有

$$\begin{aligned} \| Gx - Fx \| &= \left\| \sum_i q_i(x)(y_i - Fx) \right\| \\ &\leqslant \sum_i q_i(x) \| y_i - Fx \| \\ &\leqslant \varepsilon \sum_i q_i(x) = \varepsilon. \end{aligned}$$

直接看出 $GD \subset \operatorname{span}\{y_i\} \triangleq Z$，$Z$ 是 Y 的有限维子空间. □

§4.4.2 不动点定理

关于紧算子的最有意义的结果是不动点定理. 这一类的结果很多，其中最著名的是以下定理.

定理 4.4.5 (Schauder 不动点定理) 设 $D \subset X$ 是非空有界闭凸集，$F : D \to D$ 是紧算子，则 F 在 D 上有不动点.

若 $\dim X < \infty$，D 如定理 4.4.5，则 $F \in C(D,D)$ 必为紧算子，于是从定理 4.4.5 得到：

推论 4.4.6 设 $D \subset X$ 是非空有界闭凸集，$\dim X < \infty$，$F \in C(D,D)$，则 F 在 D 上有不动点.

特别，取 $X = \mathbf{R}^n$，$D = \bar{B}_1(0)$，从推论 4.4.6 得到著名的 Brouwer 不动点定理.

定理 4.4.7 (Brouwer 不动点定理) 设 $D = \bar{B}_1(0) \subset \mathbf{R}^n$，$F \in C(D,D)$，则 F 在 D 上有不动点.

对照定理 4.4.5 与推论 4.4.6 看出，当空间 X 从有限维过渡到无限维时，对 F 的要求从仅仅连续过渡到为紧算子. 由此可见，至少从不动点性质来看，紧算子起有限维空间中的连续算子的作用.

应当说，不动点定理 4.4.5 的表述是高度简洁清晰的. 与此形成鲜明对照的是，其证明却颇为复杂，以至不宜列入本书的任务. 不过，如果你感兴趣的话，还是可从如下的简略勾画中获得某些初步印象.

证明思路 前面已经指出，定理 4.4.5 ⇒推论 4.4.6 ⇒定理 4.4.7；而实际证明过程则恰好将这一顺序倒过来，即首先直接证明定理 4.4.7，然后由其推出推论 4.4.6，进而推出定理 4.4.5.

(i) Brouwer 不动点定理的证法是多种多样的，但本质上都基于一定的拓扑概念，因而远不是平凡的. 此处我们认可 Brouwer 定理的结论，并将它当作进一步推理的出发点.

(ii) 由定理 4.4.7 推出推论 4.4.6 的关键在于利用如下事实：

(P) 设 $\dim X < \infty$，$D \subset X$ 是非空有界闭凸集，则存在 $n \leqslant \dim X$，使得 D 同胚(参看定义 1.7.6)于 $\mathbf{R}^n$ 中的闭单位球.

现在设 D, F 如推论 4.4.6，取同胚 $G : D \to \bar{B}_1(0) \subset \mathbf{R}^n$，则

$$GFG^{-1} : \bar{B}_1(0) \to \bar{B}_1(0)$$

是一连续映射. 由定理 4.4.7，存在 $y \in \bar{B}_1(0)$，使得 $GFG^{-1}y = y$. 令 $x = G^{-1}y$，则

$$Fx = G^{-1}GFG^{-1}y = G^{-1}y = x \in D,$$

可见 x 是 F 在 D 上的不动点.

虽然利用命题(P)轻而易举地从定理 4.4.7 推出了推论 4.4.6,但命题(P)的证明却远不简单,它依赖于某些精细的构造,不是这个简短概括所能介绍的.

(iii) 从推论 4.4.6 推出定理 4.4.5. 设 D, F 如定理 4.4.5. $\forall n \in \mathbf{N}$, 由定理 4.4.4,有 $F_n \in C(D, X)$, 使得

$$\| F_n x - Fx \| < 1/n; \quad (x \in D) \tag{4.4.7}$$

$$F_n D \subset D_n \triangleq \left\{ \sum q_i y_i : q_i \geqslant 0, \sum q_i = 1 \right\},$$

其中 $\{y_i\} \subset FD$ 是一有限集. 于是 D_n 是 Y 的某个有限维子空间的有界闭凸子集, 由 $\{y_i\} \subset FD \subset D$ 及 D 为凸集推出 $D_n \subset D$, 因此 $F_n D_n \subset F_n D \subset D_n$. 对 F_n : $D_n \to D_n$ 应用推论 4.4.6 得出 $x_n = F_n x_n \in D_n (\forall n \in \mathbf{N})$. 由 F 的紧性,不妨设 $Fx_n \to x (n \to \infty)$. 由

$$\begin{aligned} \| x_n - x \| &\leqslant \| F_n x_n - F x_n \| + \| F x_n - x \| \\ &< 1/n + \| F x_n - x \| \qquad (\text{用式}(4.4.7)) \end{aligned}$$

推出 $x_n \to x \in D (n \to \infty$, 注意 D 为闭集). 于是

$$x = \lim_n F x_n = F(\lim_n x_n) = Fx,$$

可见 x 是 F 在 D 上的不动点. □

§4.4.3 某些应用

如同压缩映射原理一样,Schauder 不动点定理亦是研究方程问题的重要工具. 与压缩映射原理比较,Schauder 不动点定理有一些明显的优点,最主要的是它排除了对映射的压缩性要求,这就实质性地扩大了应用范围,特别是能应用到那些不能确立惟一性的方程问题. 当然,为获得这一优势而付出的代价也不可低估:Schauder 不动点定理本身不包含任何得出不动点的构造程序,因而它除了用于纯存在性论证之外,并无数值计算的功能. 此外,对于映射的紧性及集 D 的凸性要求,都是较强的限制. 不过,在 Schauder 不动点定理的进一步推广中,这些限制是可以放宽的.

Schauder 不动点定理的应用范围广泛,大多涉及较深入的专门知识与繁琐的技术性推演,适于用作导引性解释的简单例子并不多. 下面的几个例子无疑是比较平凡的,不过也能说明一定的问题.

例 4.4.8 仍然考虑例 4.2.4 中讨论过的初值问题

$$y'(t) = f(t, y(t)), \quad y(a) = x_0 \in X, \tag{4.4.8}$$

不过,现在设 $\dim X < \infty$ (例如 $X = \mathbf{R}^n$). 仍设 $f \in C(J \times X, X)$, $J = [a, b]$($a < b$); 但不假定 f 满足任何 Lipschitz 条件. 则存在 $\beta \in (a, b]$, 使得问题(4.4.8)式在区间 $[a, \beta]$ 上有解.

证 大体上仍循例 4.2.4 中已程序化的步骤考虑. 令 $Y = C([a, \beta], X)$, $\beta \in (a, b]$ 是待定的;在 Y 中使用范数

$$\| y \| = \sup_{a \leqslant t \leqslant \beta} \| y(t) \| \quad (y \in Y).$$

将问题(4.4.8)转化为不动点方程 $y = Fy$, 此处

$$Fy(t) = x_0 + \int_a^t f(s, y(s)) \mathrm{d}s \quad (y \in Y). \tag{4.4.9}$$

如同例 4.4.2,可证明 $F: Y \to Y$ 为紧算子①. 余下只要选择有界闭凸集 $D \subset Y$, 使得可验证 $FD \subset D$. 最简单的选择是取 D 为闭球

$$D = \bar{B}_\rho(x_0) = \{ y \in Y : \| y - x_0 \| \leqslant \rho \},$$

此处 ρ 是任意取定的正数, x_0 看作恒等于 x_0 的函数. 令

$$M = \sup\{ \| f(t, x) \| : t \in J, x \in X, \| x - x_0 \| \leqslant \rho \}.$$

由 f 的连续性推出 $M < \infty$ (注意 $\dim X < \infty$). $\forall y \in D$, 有

$$\| Fy - x_0 \| = \sup_{a \leqslant t \leqslant \beta} \left\| \int_a^t f(s, y(s)) \mathrm{d}s \right\| \qquad (\text{用式}(4.4.9))$$

$$\leqslant \int_a^\beta \| f(s, y(s)) \| \mathrm{d}s \leqslant M(\beta - a).$$

可见,只要 $a < \beta \leqslant \min\{b, a + \rho/M\}$, 就可保证 $\| Fy - x_0 \| \leqslant \rho (\forall y \in D)$, 即 $FD \subset D$. 因 D 显然是有界闭凸集,故对 F 应用定理 4.4.5 即得所要证. □

与例 4.2.4 不同,例 4.4.8 既未得出解的惟一性,也未确立在整个区间 J 上解的存在性. 实际上,排除 Lipschitz 条件(式(4.2.14))之后,以上两点都未必成立,这可用浅显的例子来说明. 例如,初值问题

$$y' = \sqrt{y}, \quad y(0) = 0$$

在区间 $[0, \infty)$ 上有两个解: $y \equiv 0$ 与 $y = x^2/4$; 初值问题

---

① 证明 F 的紧性时,需对连续函数空间 $C([a, \beta], X)$ 应用 Arzela-Ascoli 定理,该定理(定理 1.4.6)虽然是对空间 $C(J)$ 表述的,但容易说明它亦适用于 $C(J, X)$($\dim X < \infty$).

$$y' = 1 + y^2, \quad y(0) = 0$$

的解 $y = \tan t$ 定义于 $[0,\pi/2)$ 上，而不能定义于 $[0,\pi/2]$ 上.

例 4.4.9(2 阶边值问题)　考虑以下边值问题：

$$\begin{cases} x''(t) = f(t,x(t),x'(t))(t \in J = [0,1]), \\ x(0) = x(1) = 0, \end{cases} \tag{4.4.10}$$

其中 $f \in C(J \times \mathbf{R}^2, \mathbf{R})$. 在微分方程论中证明了：$x \in C^1(J)$ 是问题(4.4.10)的解当且仅当它满足积分方程

$$x(t) = -\int_0^1 G(t,s)f(s,x(s),x'(s))\mathrm{d}s, \tag{4.4.11}$$

其中

$$G(t,s) = \begin{cases} s(1-t), & 0 \leqslant s \leqslant t, \\ t(1-s), & t < s \leqslant 1. \end{cases} \tag{4.4.12}$$

于是将问题(4.4.10)的可解性归结为不动点方程 $x = Fx$ 的可解性，此处 $Fx(t)$ 记式(4.4.11)之右端. F 是定义于空间 $X = C^1(J)$ 上的算子，在 $C^1(J)$ 中采用范数(依式(1.2.6))

$$\| x \|_1 = \max\{ \| x \|_0, \| x' \|_0 \}.$$

取定 $x \in X$，令 $y(t) = f(t,x(t),x'(t))$，则 $y \in C(J)$，

$$Fx(t) = \int_0^t (t-s)y(s)\mathrm{d}s - t\int_0^1 (1-s)y(s)\mathrm{d}s;$$

$$(Fx)'(t) = \int_0^t y(s)\mathrm{d}s - \int_0^1 (1-s)y(s)\mathrm{d}s$$

$$= -\int_0^1 G_t(t,s)y(s)\mathrm{d}s. \tag{4.4.13}$$

由此可见 $Fx \in X$，因而 $F: X \to X$. 若 $\{x_n\} \subset X, x_n \rightrightarrows x, x'_n \rightrightarrows x'$，则一个平凡的论证可说明

$$Fx_n \rightrightarrows Fx, \quad (Fx_n)' \rightrightarrows (Fx)'.$$

这表明 F 连续. 设 $A \subset X$ 有界，令 $\beta = \sup_{x \in A} \| x \|_1$. 今证 $B \triangleq FA$ 在 X 中相对紧. 为此，只要证 B 与 $B' \triangleq \{x' : x \in B\}$ 在 $C(J)$ 中有界且等度连续. 令

$$M = \sup\{ | f(t,x,y) | : t \in J, x, y \in [-\beta, \beta] \}. \tag{4.4.14}$$

$\forall x \in A$，有

$$\| Fx \|_0 = \sup_{t\in J}\left|\int_0^1 G(t,s)f(s,x(s),x'(s))\mathrm{d}s\right|$$

$$\leqslant \sup_{t\in J}\int_0^1 G(t,s)\mathrm{d}s \cdot M = \frac{M}{8};\quad (\text{用式}(4.4.12)) \qquad (4.4.15)$$

$$\|(Fx)'\|_0 \leqslant \sup_{t\in J}\int_0^1 |G_t(t,s)|\mathrm{d}s \cdot M = \frac{M}{2};\quad (\text{用式}(4.4.13))$$

$$(4.4.16)$$

$$|Fx(t) - Fx(\tau)| \leqslant M\int_0^1 |G(t,s) - G(\tau,s)|\mathrm{d}s;$$

$$|(Fx)'(t) - (Fx)'(\tau)| \leqslant M|t-\tau|,$$

$(t,\tau \in J)$. 以上不等式表明 B 与 B' 有界且等度连续，故 F 为紧算子.

余下的问题是确定有界闭凸集 $D \subset X$，使得 $FD \subset D$. 例 4.4.8 的经验使我们倾向于选择 D 为闭球：

$$D = \{x \in X : \|x\|_0 \leqslant \beta,\ \|x'\|_0 \leqslant \beta\}, \qquad (4.4.17)$$

具中 β 是待定正数. 若 M 依(14)(但(14)中的 β 依此处的意义)，则由式(4.4.15)，式(4.4.16)看出，要 $FD \subset D$，必须 $M \leqslant 2\beta$. 如果没有关于 f 的附加假设，此条件未必能满足. 现在设 f 是线性增长的，即

$$|f(t,x,y)| \leqslant k(1+|x|+|y|) \quad (t \in J, x, y \in \mathbf{R}), \qquad (4.4.18)$$

则由式(4.4.14)有 $M \leqslant k(1+2\beta)$. 于是不等式 $M \leqslant 2\beta$ 由

$$k(1+2\beta) \leqslant 2\beta$$

推出. 若 $0 < k < 1$，则只要取 $\beta \geqslant k/2(1-k)$，然后定义 D 如式(4.4.17)，就可保证 $FD \subset D$. 因此，当条件(4.4.18)满足且 $0 < k < 1$ 时，问题(4.4.10)有解.

例 4.4.10(周期轨道)　设 $D \subset \mathbf{R}^n$，$F \in C(D,D)$. 则 F 在 D 上定义一离散动力系统. $\forall x_0 \in D$，由

$$x_k = Fx_{k-1} = F^k x_0, \quad (k \geqslant 1)$$

(参照式(4.2.9))定义的序列 $\{x_k : k \geqslant 0\}$ 称为系统从点 x_0 发出的轨道. 若存在 $p \geqslant 1$，使得 $F^p x_0 = x_0$，则称 x_0 为 F 的以 p 为周期的**周期点**，称从 x_0 发出的轨道为周期轨道，它至多由 p 个点($x_0, Fx_0, \cdots, F^{p-1}x_0$)组成. 显然 $p \geqslant 1$ 是 F 的一个周期相当于 F^p 有不动点. 因此，周期轨道的存在性就转化成了一个不动点问

题.而由定理4.4.6,要验证 $p\geqslant 1$ 是 F 的周期,只需求得非空有界闭凸集 $B\subset D$,使得 $F^pB\subset B$.这件事可能并不简单,但通常比直接寻求周期点要容易些.

§4.5 凸函数

在非线性函数中,凸函数占有特别重要的地位.凸函数既有较好的性质(与非凸函数比较),又有较大的适用性(与线性函数比较),因而在现代数学及其应用中被广泛使用.近年来,凸分析及与之关联的最优化理论的深入发展与广泛应用,大大激发了对凸函数的研究兴趣.在力图将泛函分析的标准内容与适当的应用领域联系起来的本书中,凸函数自然占有一席之地.

§4.5.1 凸函数及其性质

以下设 $D\subset X$ 是给定的非空凸集.

定义 4.5.1 设 $f:D\to\mathbf{R}$.若 $\forall x,y\in D,t\in(0,1)$,$f$ 满足不等式

$$f(t'x+ty)\leqslant t'f(x)+tf(y) \tag{4.5.1}$$

(此处及以下都约定 $t'=1-t$),则称 f 为**凸函数**;若当 $x\neq y$ 时式(4.5.1)为严格不等式,则称 f 为**严格凸函数**.当 $-f$ 是凸(或严格凸)函数时,称 f 为**凹**(或严格凹)函数.

用归纳法,可将条件(1)推广为一般的 **Jensen 不等式**

$$f\Big(\sum_i t_ix_i\Big)\leqslant\sum_i t_if(x_i), \tag{4.5.2}$$

其中 $\sum t_ix_i$ 是 D 中点的任一凸组合(参考§2.4.5).注意到凸组合原不过是一种加权平均,因而不等式(4.5.2)可解释为:f 在个别点的值的加权平均比在混合情况 $\sum t_ix_i$ 下更具优势;凹函数则恰好相反.凡对状态的多样性表现出一定倾向的领域,例如生态与经济问题,常不免要用到凸(或凹)函数.

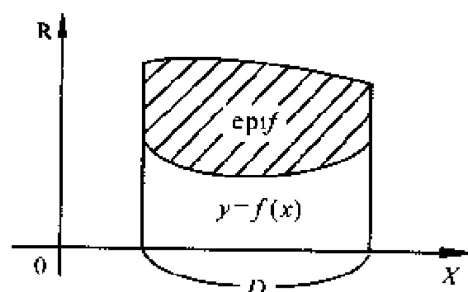


图 4-1

凸函数有明显的几何解释.令

$$\mathrm{epi}f=\{(x,r)\in D\times\mathbf{R}:f(x)\leqslant r\}, \tag{4.5.3}$$

称它为 f 的**上图**(见图 4-1).容易验证:f 是凸函数 $\Leftrightarrow$ epi$f$ 是 $X\times\mathbf{R}$ 中的凸集.顺便指出 epi$f$ 的另一个性质:若 D 是闭集,则 epi$f$ 是 $X\times\mathbf{R}$ 中的闭集 $\Leftrightarrow f$ **下半连续**,

这意味着

$$x_n \to x \Rightarrow f(x) \leqslant \underline{\lim}_n f(x_n) \quad (\{x_n\} \subset D). \tag{4.5.4}$$

事实上,若 epi$f$ 是闭集, $\{x_n\} \subset D, x_n \to x$, 则由 $(x_n, f(x_n)) \in$ epi$f$ 推出

$$(x, \underline{\lim}_n f(x_n)) \in \text{epi} f,$$

从而 $f(x) \leqslant \underline{\lim}_n f(x_n)$. 反之,若条件(4.5.4)满足, $\{x_n\} \subset D, f(x_n) \leqslant r_n, (x_n, r_n) \to (x, r)$, 即 $x_n \to x, r_n \to r (n \to \infty)$, 则由条件(4.5.4)有

$$f(x) \leqslant \underline{\lim}_n f(x_n) \leqslant \lim_n r_n = r,$$

这表明 $(x, r) \in$ epi$f$, 因此 epi$f$ 是闭集.

凸函数与凸集的另一关系是:若 f 是凸函数,则对任给 $r \in \mathbf{R}$, 集 $\{f \leqslant r\}$ 是凸集: $\forall x, y \in \{f \leqslant r\}, \forall t \in (0,1)$, 有

$$f(t'x + ty) \leqslant t'f(x) + tf(y) \leqslant (t' + t)r = r,$$

从而 $t'x + ty \in \{f \leqslant r\}$. 类似地,当 f 为凸函数时 $\{f < r\}$ 亦为凸集. 以上结论为构成凸集提供了一个一般方法. 容易注意到,这恰好与用连续函数的性质构作开集或闭集类比(参看定理 1.3.6(ii)). 不过应注意, $\{f \leqslant r\}$ 恒为凸集并不是 f 为凸函数的充分条件.

现在看一些凸函数的例子.

例 4.5.2 (i) 设 $A \subset X$ 是非空凸集, $f(x) = d(x, A)$. 任给 $t \in (0,1)$, $x, y \in X, a, b \in A$, 有

$$\begin{aligned} f(t'x + ty) &\leqslant \| t'x + ty - (t'a + tb) \| \\ &\leqslant t' \| x - a \| + t \| y - b \|. \end{aligned}$$

固定 t, x, y, 令 a, b 在 A 中变动,使 $\| x - a \| \to f(x)$, $\| y - b \| \to f(y)$, 即得不等式(4.5.1). 因此, $f: X \to \mathbf{R}$ 是凸函数. 由此又推出, $\forall r > 0$, 集

$$\{x \in X : d(x, A) < r\} \text{ 与 } \{x \in X : d(x, A) \leqslant r\}$$

分别为开凸集与闭凸集. 你大概还记得,我们在分离定理(定理 2.4.14)的证明中曾用过这些凸集.

特别取 $A = \{x_0\}$, 得出 $f(x) = \| x - x_0 \|$ 是凸函数.

(ii) 设 H 是 Hilbert 空间, A 是 H 上的正自伴算子(参看定义 3.4.3 与定义 3.4.7),则二次泛函 $f(x) = \langle Ax, x \rangle$ 是凸函数: $\forall t \in (0,1), x, y \in H$, 有

$$f(t'x + ty) = t'^2 f(x) + tt'\langle Ax, y \rangle + tt'\langle Ay, x \rangle + t^2 f(y)$$

$$\leqslant t'f(x) + tf(y),$$

后一不等号依据不等式

$$\langle Ax, y\rangle + \langle Ay, x\rangle \leqslant f(x) + f(y),$$

而上式由不等式 $\langle A(x-y), x-y\rangle \geqslant 0$ 推出.

以上结论可看作如下事实的推广:设 $a \geqslant 0$, 则二次函数 $ax^2(x \in \mathbf{R})$ 是凸函数.

关于凸函数的运算,容易验证如下简单结论:凸函数的非负系数线性组合为凸函数;若 f 为凸函数, g 是实区间上的单调增凸函数,则 $g(f(x))$ 是凸函数;若 $f(x)$ 是非负凸函数,则 $f^2(x)$ 是凸函数;特别, $f(x) = \|x\|^2$ 是凸函数.

更深入一点的性质是连续性与可微性.基本的结论是:凸函数远比其他函数更容易成为连续函数与可微函数.你不妨想像一下:一条凸的曲线容易出现间断与不光滑吗?

定理 4.5.3　设 $D \subset X$ 是凸开集, $f: D \to \mathbf{R}$ 是凸函数.则 f 在 D 内连续 $\Leftrightarrow f$ 在 D 内某点 x_0 邻近有上界;若 $\dim X < \infty$, 则 f 必连续.

证　设 $x_0 \in D$, f 在某个球 $B_r(x_0)$ 内有上界 β. 令 $\delta = r/2$, 设 $x \in B_\delta(x_0)$, $\rho = \|x - x_0\| > 0$, 则 $z \triangleq x + \rho^{-1}\delta(x - x_0) \in B_r(x_0)$,

$$f(x) = f\left(\frac{\rho z + \delta x_0}{\rho + \delta}\right) \leqslant \frac{\rho f(z) + \delta f(x_0)}{\rho + \delta};\text{(用式(4.5.1))}$$

$$f(x) - f(x_0) \leqslant \frac{\rho[f(z) - f(x_0)]}{\rho + \delta}$$

$$\leqslant \frac{\beta - f(x_0)}{\delta}\|x - x_0\|.$$

类似地,由 $y \triangleq x_0 + \rho^{-1}\delta(x_0 - x) \in B_\delta(x_0)$ 推出

$$f(x_0) - f(x) = f\left(\frac{\rho y + \delta x}{\rho + \delta}\right) - f(x) \leqslant \frac{\rho[f(y) - f(x)]}{\rho + \delta}$$

$$= \frac{\rho\{[f(y) - f(x_0)] + [f(x_0) - f(x)]\}}{\rho + \delta},$$

由此解出

$$f(x_0) - f(x) \leqslant \frac{\rho}{\delta}[f(y) - f(x_0)]$$

$$\leqslant \frac{\beta - f(x_0)}{\delta}\|x - x_0\|.$$

这表明 f 在 x_0 连续.

为证 f 在 D 内连续,只需指明: $\forall x \in D$, f 在 x 邻近有上界.取定 $x \in D$, $k > 1$, 使 $z \triangleq x_0 + k(x - x_0) \in D$, 则当 y 邻近 x 时 $\frac{ky - z}{k - 1}$ 邻近 x_0, 于是

$$\begin{aligned} f(y) &= f\left(\frac{z}{k} + \frac{k-1}{k}\frac{ky - z}{k - 1}\right) \\ &\leqslant \frac{1}{k}f(z) + \frac{k-1}{k}f\left(\frac{ky - z}{k - 1}\right) \quad (\text{用式}(4.5.1)) \\ &\leqslant \frac{1}{k}f(z) + \frac{k-1}{k}\beta, \end{aligned}$$

可见 f 在 x 邻近有上界.

若 $\dim X = n < \infty$, 则可在 D 内取出 $x_i (0 \leqslant i \leqslant n)$, 使得 $\{x_i - x_0 : 1 \leqslant i \leqslant n\}$ 线性无关,从而以 $x_i (0 \leqslant i \leqslant n)$ 为顶点的多面体①含内点 $\bar{x}$. 用 Jensen 不等式(4.5.2)显然推出 f 在 A 内有上界,从而在 $\bar{x}$ 邻近有上界,于是用前段已证结论得出 f 连续. □

有限维空间中凸开集上的凸函数必连续,这无疑是一个极有意义的结论.但你在使用这一结论时不可越出其适用范围.首先,在有限维空间中,非开集上的凸函数未必连续(当然间断只能出现在边界上),用如下函数

$$f(x) = \begin{cases} 0, & 0 \leqslant x < 1, \\ 1, & x = 1 \end{cases}$$

即可说明这一点.也不要指望无限维空间上的凸函数一定连续,用无界线性泛函即可说明这一点,线性泛函显然是凸函数(同时也是凹函数).

定理 4.5.4 设 $f : D \to \mathbf{R}$ 是凸函数, $x \in D$.

(i) 任给 $h \in X$, 若 $x \pm h \in D$, 则单侧方向导数 $D_+ f(x, h)$ (参看定义 4.1.8)必存在,且

$$D_- f(x,h) \leqslant D_+ f(x,h) \leqslant \Delta f(x,h), \tag{4.5.5}$$

其中 $\Delta f(x,h) = f(x + h) - f(x)$ (下文中保持此约定).

(ii) 若 $\dim X < \infty$, $x \in D^\circ$, G 导数 $Df(x)$ 存在,则 $f'(x) = Df(x)$. 因此,对于有限维空间中的凸函数, G 可微与 F 可微一致,且梯度 $\nabla f(x) = f'(x) = Df(x)$ 有确定的意义.

---

① 以 $x_i (0 \leqslant i \leqslant n)$ 为顶点的多面体定义为点组 $\{x_i\}$ 的凸组合之全体.

证 (i) 取定 $h \in X$，令 $\varphi(t) = t^{-1}\Delta f(x, th)$，则 $\varphi(1) = \Delta f(x, h)$；当 $\varphi(0^+)$ 存在时 $D_\pm f(x, h) = \varphi(0^\pm)$，而不等式(4.5.5)则相当于 $\varphi(0^-) \leqslant \varphi(0^+) \leqslant \varphi(1)$. 因此，只要证 $\varphi(t)$ 在 $t < 0$ 时单调减，在 $t > 0$ 时单调增，且 $s < 0 < t \Rightarrow \varphi(s) \leqslant \varphi(t)$. 这三者的证明都基于不等式(4.5.1)，不妨只证 $\varphi(t)$ 在 $t > 0$ 时单调增. 设 $0 < t < s$，则

$$f(x + th) = f\left(\left(1 - \frac{t}{s}\right)x + \frac{t}{s}(x + sh)\right)$$

$$\leqslant \left(1 - \frac{t}{s}\right)f(x) + \frac{t}{s}f(x + sh), \quad (\text{用式}(4.5.1))$$

这推出 $\varphi(t) \leqslant \varphi(s)$.

(ii) 不妨设 $X = \mathbf{R}^n$，$\{e_i\}$ 是其标准基. 令 $g(h) = \Delta f(x, h) - Df(x)h$，则 g 是凸函数；由式(4.5.5)推出 $g(h) \geqslant 0$；由 G 导数的定义 4.1.8 推出 $g(th) = o(t)(t \to 0)$. 于是

$$g(h) = g\left(\sum_i h_i e_i\right) \leqslant \frac{1}{n}\sum_i g(nh_i e_i) \qquad (\text{用式}(4.5.2))$$

$$= \sum_{h_i \neq 0} h_i \frac{g(nh_i e_i)}{nh_i}$$

$$\leqslant \| h \| \left\{\sum_{h_i \neq 0}\left[\frac{g(nh_i e_i)}{nh_i}\right]^2\right\}^{1/2} = o(\| h \|),$$

这表明 $f'(x)h = Df(x)h$，即 $f'(x) = Df(x)$. □

更深入的分析可以指明：有限维空间上的凸函数必几乎处处可微. 从无界线性泛函的例子看出，无限维空间上的凸函数可以处处不可微. 由此可见，对于凸函数的连续性与可微性，有限维空间与无限维空间的差别甚大.

§4.5.2 次微分

即使在有限维空间上，凸函数也未必处处可微，函数 $f(x) = |x|$（它在 $x = 0$ 不可微）提供了最简单的例证. 对于那些需用导数表述条件的问题（如极值问题），即使个别点不可微，亦会带来很大的不便. 在这种情况下，人们用某种广义的微分来代替导数. 次微分正是这样一种广义微分.

定义 4.5.5 设 $f: D \to \mathbf{R}$，$x \in D$，令

$$\partial f(x) = \{u \in X^*: u(h) \leqslant \Delta f(x, h)(\forall h \in X)\} \qquad (4.5.6)$$

（当 $z \notin D$ 时约定 $f(z) = \infty$，下面概如此）. 若 $\partial f(x) \neq \varnothing$，则说 f 在 x **次可微**，

称 $\partial f(x)$ 为 f 在 x 的**次微分**,称任何 $u \in \partial f(x)$ 为 f 在 x 的**次梯度**.

注意定义 4.5.5 中并未假定 f 为凸函数.不过,若 f 在 D 上处处次可微,则可推出 f 为凸函数.因此可以说,次微分概念实际上是专为凸函数而设的.

次微分有很直观的几何解释.设 $u \in \partial f(x_0)$,令 $v = (-u, 1) \in X^* \times \mathbf{R} = (X \times \mathbf{R})^*$ ①,则不等式

$$u(x - x_0) \leqslant f(x) - f(x_0) \quad (\forall x \in D) \tag{4.5.7}$$

相当于

$$v(x, f(x)) \geqslant v(x_0, f(x_0)) \triangleq c \quad (\forall x \in D). \tag{4.5.8}$$

不等式(4.5.8)意味着,在空间 $X \times \mathbf{R}$ 中,f 的图形 $G(f) = \{(x, f(x)) : x \in D\}$ 完全位于超平面 $v(x, y) = c$ 的一侧;或者说,超平面 $v(x, y) = c$ 在点 $(x_0, f(x_0))$ 支承起 $G(f)$,当然也支承起上图 epi$f$(参看图 4-2).因 f 在 x_0 的次梯度未必是惟一的,故上述的支承超平面也未必是惟一的.这与 f 在其可微点有惟一切平面形成对照.

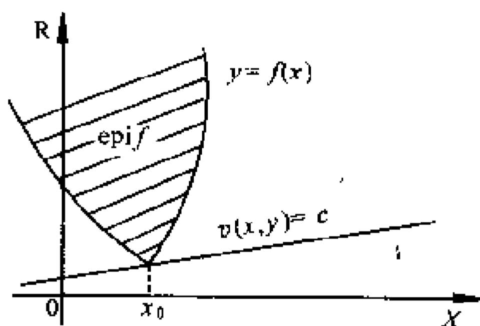


图 4-2

现在看两个具体例子.

例 4.5.6 (i) 设 $f : \mathbf{R} \to \mathbf{R}$ 是凸函数.由定理 4.5.3,f 处处连续;由定理 4.5.4,$f'_{\pm}(x)$ 处处存在,且 $f'_{-}(x) \leqslant f'_{+}(x)$.取定 $x \in \mathbf{R}$,$\forall u \in \mathbf{R}(= \mathbf{R}^*)$,由式(4.5.6)有

$$\begin{aligned} u \in \partial f(x) &\Leftrightarrow uh \leqslant \Delta f(x, h) \quad (\forall h \in \mathbf{R}), \\ &\Leftrightarrow \frac{\Delta f(x, -h)}{-h} \leqslant u \leqslant \frac{\Delta f(x, h)}{h} \quad (\forall h > 0) \\ &\Leftrightarrow f'_{-}(x) \leqslant u \leqslant f'_{+}(x), \end{aligned}$$

故得

$$\partial f(x) = [f'_{-}(x), f'_{+}(x)] \subset \mathbf{R}. \tag{4.5.9}$$

因此,$f'(x)$ 存在 $\Leftrightarrow f'_{-}(x) = f'_{+}(x) \Leftrightarrow \partial f(x)$ 只含一点,即 $f'(x)$.

① 对任何赋范空间 X、Y,用§2.3 中的方法容易建立以下一般结论:$(X \times Y)^* \cong X^* \times Y^*$,这意味着 $w \in (X \times Y)^*$ 有通式 $w(x, y) = u(x) + v(y)$,$(u, v) \in X^* \times Y^*$ 由 w 惟一决定.

(ii) 设 $f(x)=\|x\|(x\in X)$. 首先设 $0\neq x\in X$. 若 $u\in\partial f(x)$, 则

$$u(h)\leqslant\|x+h\|-\|x\|\qquad(\forall h\in X);\qquad(4.5.10)$$

$$u(x)\leqslant\|x+x\|-\|x\|=\|x\|;\qquad(\text{用式}(4.5.10))$$

$$-u(x)=u(-x)\leqslant\|x-x\|-\|x\|=-\|x\|.$$

这推出 $\|x\|=u(x)\leqslant\|u\|\|x\|$, 故 $\|u\|\geqslant 1$. 另一方面,

$$\|u\|=\sup_{\|h\|=1}u(h)\leqslant\sup_{\|h\|=1}\|x+h\|-\|x\|\quad(\text{用式}(4.5.10))$$

$$\leqslant\sup_{\|h\|=1}\|x+h-x\|=1,$$

故得 $\|u\|=1$. 反之, 若 $u\in X^*$, $u(x)=\|x\|$, $\|u\|=1$, 则式(4.5.10)必满足,因而 $u\in\partial f(x)$. 这就得到

$$\partial f(x)=\{u\in X^*:u(x)=\|x\|,\|u\|=1\}.$$

其次,直接由式(4.5.6)得出, $\forall u\in X^*$, 有

$$u\in\partial f(0)\Leftrightarrow u(h)\leqslant\|h\|(\forall h\in X)\Leftrightarrow\|u\|\leqslant 1.$$

综合以上两种情况得到

$$\partial\|x\|=\begin{cases}\{u\in X^*:u(x)=\|x\|=\|u\|\|x\|\}, & x\neq 0,\\ \overline{B}_1(0)\subset X^*, & x=0.\end{cases}\qquad(4.5.11)$$

利用式(4.5.11)可以验证:若 X 是实 Hilbert 空间, $0\neq x\in X$, 则

$$\partial\|x\|=\{x/\|x\|\}=\{\nabla\|x\|\},$$

这与式(4.1.3)恰相对应. 特别,取 $X=\mathbf{R}$ 得

$$\partial|x|=\begin{cases}\{\operatorname{sgn}x\}, & x\neq 0,\\ [-1,1], & x=0.\end{cases}\qquad(4.5.12)$$

对于 $f(x)=|x|$, 等式 $\partial f(0)=[-1,1]$ 在几何上意味着 $\forall k\in[-1,1]$, $\mathbf{R}^2$ 中的直线 $y=kx$ 在原点处支承起函数 $f(x)=|x|$ 的图形;若 $|k|>1$, 则直线 $y=kx$ 不再有上述性质.

下面的定理表明,次可微是一个很弱的条件.

命题 4.5.7 设 $f:D\to\mathbf{R}$ 是凸函数, $x\in D$.

(i) 必要条件:若 f 在 x 次可微,则 f 在 x 下半连续.

(ii) 充分条件:若 f 在 x 连续,则 f 在 x 次可微.

总而言之,对于凸函数次可微介于下半连续与连续之间.

证 (i) 设 $u \in \partial f(x)$,则当 $\{x_n\} \subset D, x_n \to x(n \to \infty)$ 时有

$$f(x) \leqslant \varliminf_n [f(x_n) - u(x_n - x)] = \varliminf_n f(x_n),$$

可见 f 在 x 下半连续(依式(4.5.4)).

(ii) $\forall \varepsilon > 0$, f 在 x 连续推出有 x 的邻域 U,使得

$$f(y) < f(x) + \varepsilon \triangleq \rho - \varepsilon \quad (\forall y \in U).$$

这表明 $U \times (\rho - \varepsilon, \infty) \subset \mathrm{epi} f$,因而 (x, ρ) 是 $\mathrm{epi} f$ 的内点. 在积空间 $X \times \mathbf{R}$ 中对凸集 $A = \mathrm{epi} f$ 与 $B = \{(x, f(x))\}$ 应用分离定理 2.4.14(i),得出 $v \in (X \times \mathbf{R})^*$,使得

$$v(A^\circ) < v(B), \quad v(A) \leqslant v(B).$$

这意味着有 $u \in X^*$ 与 $\beta \in \mathbf{R}$,使得

$$\begin{cases} u(x) + \beta\rho < u(x) + \beta f(x), & (4.5.13a) \\ f(y) \leqslant r \to u(y) + \beta r \leqslant u(x) + \beta f(x). & (4.5.13b) \end{cases}$$

由 $\rho = f(x) + 2\varepsilon > f(x)$ 与式(4.5.13a)得出 $\beta < 0$;在式(4.5.13b)中取 $r = f(y)$ 得

$$-\beta^{-1} u(y - x) \leqslant f(y) - f(x) \quad (\forall y \in D),$$

这表明 $-\beta^{-1} u \in \partial f(x)$,故 f 在 x 次可微. □

在引进次微分时我们已指出,它将用作导数的某种替代. 而例 4.5.6 中的例子似乎表明,当导数存在时恰与次微分一致. 一个自然的问题是:次微分与导数之间,一般地具有什么关系? 下面的定理对此给出完全的回答.

定理 4.5.8 设 $f: D \to \mathbf{R}$ 是凸函数, $x \in D$.

(i) 设 $u \in X^*$,则 $u \in \partial f(x) \Leftrightarrow u(h) \leqslant D_+ f(x,h) \Leftrightarrow u(h) \geqslant D_- f(x,h) (\forall h \in X)$.

(ii) 若 G 导数 $Df(x)$ 存在,则 $\partial f(x) = \{Df(x)\}$.

(iii) 若 f 在 x 连续, $\partial f(x) = \{u\}$,则 $u = Df(x)$. 因此,对于连续凸函数 $f(x)$, $Df(x)$ 存在 $\Leftrightarrow \partial f(x)$ 只含一个元素;对于 $\mathbf{R}^n$ 上的凸函数 $f(x)$, $f'(x)$ 存在 $\Leftrightarrow \partial f(x)$ 只含一个元素.

证 (i) 由式(4.5.5)与 $D_- f(x,h) = -D_+ f(x,-h)$ (式(4.1.20))推出.

(ii) 由(i)推出.

(iii) 取定 $0 \neq h \in X, \alpha \in [D_- f(x,h), D_+ f(x,h)]$. 定义

$$u_\alpha : \mathbf{R}h \to \mathbf{R}, \quad th \mapsto \alpha t,$$

则 u_α 是 X 的 1 维子空间 $\mathbf{R}h$ 上的线性泛函，由 α 的选择知

$$u_\alpha(th) \leqslant \mathrm{D}_+ f(x, th).$$

可验证 $\mathrm{D}_+ f(x,h)$ 对 h 是次线性的（这要用到 f 的凸性），于是由 Hahn-Banach 定理 2.4.5(i)推出：u_α 可延拓为 X 上的线性泛函（仍记作 u_α），使得

$$u_\alpha(z) \leqslant \mathrm{D}_+ f(x, z) \leqslant \Delta f(x, z) \quad (\forall z \in X).$$

由 f 在 x 连续推出，当 z 接近于 0 时 $\Delta f(x,z)$ 有界，因而 $u_\alpha(z)$ 在 $z = 0$ 邻近有界，故 $u_\alpha \in X^*$. 由 $u_\alpha(z) \leqslant \Delta f(x,z)(\forall z \in X)$ 得 $u_\alpha \in \partial f(x)$. 但 $\partial f(x) = \{u\}$，故 $u_\alpha = u$ 与 α 无关. 这又推出 $\mathrm{D}_- f(x,h) = \mathrm{D}_+ f(x,h) = u(h)$，因此 $\mathrm{D}f(x) = u$. □

§4.6 极值理论

关于极值，你并不陌生，在微积分学中已接触过了. 极值理论之所以重要，其根本理由在于：人们确信，无论自然与社会，任何事物在趋于平衡时均不可避免地遵循极值原则. 例如，力学中的能量最小原理；经济学中的效用最大化原则等. 简单的极值问题利用通常的微分法即可解决. 现代极值理论当然极大地深化了. 但极值问题的提法却仍如其原型一样简单明确；而极值条件则深深地植根于解决初等极值问题的那些经典命题中. 本节的目的在于让你看到，利用泛函分析的概念与方法，可将熟知的极值结论推广到何等一般的地步. 这些推广的一部分，在经典变分学中早已在稍不同的形式下给出并广泛应用了. 如我们在 §4.1 中已提到的，正是变分学的发展，在一定程度上促成了泛函分析的诞生. 极值理论对于泛函分析本身的重要性，于此可见.

§4.6.1 一般概念

给定实泛函 $f : D \subset X \to \mathbf{R}$，设 $x_0 \in D$. 下面的定义与微分学中的极值定义并无实质区别，所不同的是所用的空间框架更为一般罢了.

定义 4.6.1 若存在 $\delta > 0$，使得

$$f(x_0) \leqslant f(x) \quad (\forall x \in D \cap B_\delta(x_0)), \tag{4.6.1}$$

则称 x_0 为 f 在 D 上的**局部极小点**，称 $f(x_0)$ 为**局部极小值**. 若对 $x \neq x_0$，式(4.6.1)是严格不等式，则称 x_0 为**严格局部极小点**. 若

$$f(x_0) \leqslant f(x) \quad (\forall x \in D), \tag{4.6.2}$$

则称 x_0 为 f 在 D 上的**全局极小点**或简称为最小值点.

类似地,可定义极大点、极大值等.极小点与极大点合称为**极值点**;极小值与极大值合称为**极值**.因 f 的极大值正是 $-f$ 的极小值,故仅需展开有关极小值的理论.

若 x_0 是 f 在 D 上的局部(或全局)极小点,则通常称 x_0 为**最小化问题**

$$\min f(x), \quad x \in D \tag{4.6.3}$$

的局部(或全局)**最优解**或最优点;相应地,称 $f(x_0)$ 为问题(3)的**最优值**.无论最小化问题或最大化问题,都可统称为**最优化问题**.关于最优化问题的基本课题是:

(A) 确定(局部或全局)最优解的存在性.

(B) 在最优解存在的情况下,运用各种方法求出(通常是近似地)最优解.

为判定极值存在,现代极值理论中发展了一系列复杂的方法,那些归根结底基于紧性的方法,实际上可追溯到很简单的 Weierstrass 定理:闭区间上的连续函数必达到最大值与最小值.定理1.4.4(ii)已将 Weierstrass 定理推广到了紧集上的连续函数.下面的定理将使你看到,沿这一方向的推广还可以走多远.

定理 4.6.2 设 f 满足条件

$$\lim_{x \in D, \|x\| \to \infty} f(x) = \infty, \tag{4.6.4}$$

当 D 有界时认定条件(4.6.4)自动满足.则附加以下两条件之一可使 f 在 D 上取得最小值:

(i) X 是自反空间, D 是闭凸集, f 是下半连续的凸函数.

(ii) D 为闭集且含于 X 的某个有限维子空间, f 下半连续(依式(4.5.4)).

证 令 $\alpha = \inf_{x \in D} f(x)$, 则必可取出序列 $\{x_n\} \subset D$, 使得 $f(x_n) \to \alpha (n \to \infty)$. 条件(4.5.4)保证了 $\{x_n\}$ 有界.证明的关键在于从 $\{x_n\}$ 中取出某种收敛子列,其极限即为 f 的最小值点,收敛子列的存在依赖于条件(i)或(ii).

首先设条件(i)满足,则依定理2.5.4不妨设 $x_n \rightharpoonup x_0$, 必定 $x_0 \in D$(参看第二章习题32).今证 $f(x_0) = \alpha$ (从而 x_0 为最小值点),为此只要证 $\forall \beta > \alpha$, 有 $f(x_0) \leqslant \beta$. 令 $A = \{x \in D : f(x) \leqslant \beta\}$. 由 f 为凸函数直接推出 A 为凸集;由 f 下半连续(参看式(4.5.4))及 D 为闭集易推出 A 为闭集.由 $f(x_n) \to \alpha < \beta$, 不妨设 $f(x_n) < \beta$, 因而 $\{x_n\} \subset A$. 于是从 $x_n \rightharpoonup x_0$ 推出 $x_0 \in A$ (再次用第二章习题32),即 $f(x_0) \leqslant \beta$, 如所要证.

其次设条件(ii)满足.由定理1.4.2,不妨设 $x_n \to x_0 \in D$, 于是由 f 下半连续得

$$f(x_0)\leqslant \varliminf_n f(x_n) = \alpha \leqslant f(x_0),$$

这推出 $f(x_0) = \alpha$，故同样得所要证. □

以上定理证明中所用到的序列 $\{x_n\}$ 称为**极小化序列**.在有关极值存在性的论证中运用极小(或极大)化序列,乃是一种普遍方法.你不妨回顾一下定理1.4.4(ii),1.5.6(i)及2.5.6的证明,从中会发现某些共同的思路.不过应当指出,这些证明中所用到的极小化序列,其存在性只是在理论上被确立了;至于如何实际构作它,则完全是另一个问题,定理的证明并未给出任何解决的途径.

对于极值点,除了解决其存在性这一首要问题之外,有时也需要探讨极值点的惟一性或极值点集的结构.在这方面可以应用如下简单结果.

命题 4.6.3　设 $D\subset X$ 是凸集，$f: D\to \mathbf{R}$ 为凸函数,则有以下结论成立:

(i) 若 $x_0\in D$ 是 f 在 D 上的局部极小点,则它亦为全局极小点.

(ii) f 在 D 上的极小点之全体构成一凸集,因而不可能存在孤立的极小点,除非极小点是惟一的.

(iii) 若 f 严格凸,则 f 在 D 上至多有一个极小点.

证　(i) 任给 $x\in D$，令 $x_t = t'x_0 + tx, 0<t<1, t' = 1-t$，则由 D 为凸集有 $x_t\in D$，且当 $t\to 0$ 时 $x_t\to x_0$. 于是当 t 充分小时有

$$f(x_0)\leqslant f(x_t)\leqslant t'f(x_0) + tf(x),$$

由此推出 $f(x_0)\leqslant f(x)$. 故 x_0 是最小值点.

(ii) 令 $\alpha = \inf\limits_{x\in D} f(x)$，则 f 在 D 上的极小点之全体即集 $\{f\leqslant\alpha\}$，它是一凸集(参看§4.5.1).

(iii) 若 f 在 D 上有两个不同的极小点 $x_0, x_1, x_t = t'x_0 + tx_1, 0<t<1$，则

$$f(x_t) < t'f(x_0) + tf(x_1) = f(x_0),$$

这与 x_0 是极小点相矛盾.因此 f 至多有一个极小点. □

§4.6.2　微分条件

在微分学中有以下熟知结论:设 $f(x)$ 是区间 (a,b) 内的实函数.若 $x_0\in(a,b)$ 是 f 的极值点且 f 在 x_0 可微,则必 $f'(x_0) = 0$；若 $f'(x_0) = 0, f''(x_0) > 0$，则 x_0 是 f 的极小点.下面给出以上结论的多种形式的推广.以下结果是最直接的推广.

定理 4.6.4　设 $f: D\subset X\to\mathbf{R}$，$x_0$ 是 D 的内点.

(i) 必要条件:若 x_0 是 f 的局部极小点，$f'(x_0)$ 存在,则 $f'(x_0) = 0$.

(ii) 充分条件:若 $f\in C^2, f'(x_0) = 0, f''(x_0)$ 满足**正定性条件**:

$$\exists k > 0,\quad \forall h \in X,\quad 有\ f''(x_0)h^2 \geqslant k\|h\|^2, \tag{4.6.5}$$

则 x_0 是 f 的严格局部极小点.

证 (i) 只要证 $f'(x_0)h = 0(\forall h \in X)$. 取定 $h \in X$, 令 $\varphi(t) = f(x_0 + th)$, 则 $t = 0$ 为 $\varphi(t)$ 的局部极小点, 于是由微分学中的结论有

$$0 = \varphi'(0) = f'(x_0)h.$$

(ii) 因 $f''(x)$ 连续, 故 $\exists \delta > 0, \forall z \in B_\delta(0)$, 有

$$\|f''(x_0 + z) - f''(x_0)\| < k/2, \tag{4.6.6}$$

其中 k 依条件(4.6.5). 设 $0 \neq h \in B_\delta(0)$, 则对 $\varphi(t) \triangleq f(x_0 + th)$ 用 Taylor 公式得

$$\begin{aligned}
\Delta f(x_0, h) &= \varphi(1) - \varphi(0) - \frac{1}{2}\varphi''(\theta) \quad (0 < \theta < 1, 注意\ \varphi'(0) = 0) \\
&= \frac{1}{2}f''(x_0 + \theta h)h^2 \\
&= \frac{1}{2}f''(x_0)h^2 - \frac{1}{2}[f''(x_0) - f''(x_0 + \theta h)]h^2 \\
&\geqslant \frac{k}{2}\|h\|^2 - \frac{k}{4}\|h\|^2 \qquad (用式(4.6.5),(4.6.6)) \\
&= (k/4)\|h\|^2 > 0,
\end{aligned}$$

可见 f 在 x_0 取严格局部极小值. □

若 $f: D \subset \mathbf{R}^n \to \mathbf{R}$ 是 C^2 函数, 则

$$\begin{aligned}
f''(x)h^2 &= \frac{\mathrm{d}^2}{\mathrm{d}t^2}f(x + th)\Big|_{t=0} \\
&= \left[\frac{\mathrm{d}}{\mathrm{d}t}f'(x + th)h\right]\Big|_{t=0} \\
&= \left[\frac{\mathrm{d}}{\mathrm{d}t}\sum_i \frac{\partial f(x + th)}{\partial x_i}h_i\right]\Big|_{t=0} \\
&= \sum_{i,j} \frac{\partial^2 f(x)}{\partial x_i \partial x_j}h_i h_j,
\end{aligned}$$

其中 $h = (h_i) \in \mathbf{R}^n$. 因此, $f''(x_0)h^2$ 是以 $H \triangleq [\partial^2 f(x_0)/\partial x_i \partial x_j]$ 为其矩阵的二次型, 而条件(4.6.5)相当于 H 是正定矩阵. H 称为 f 在 x_0 的 **Hesse 矩阵**. 因此, 定

理 4.6.4(ii)包含了如下有限维结果:若 $f: D \subset \mathbf{R}^n \to \mathbf{R}$ 是 C^2 函数, $\nabla f(x_0) = 0$, f 在 x_0 的 Hesse 矩阵正定,则 x_0 是 f 的严格局部极小点.

从形式上看,定理 4.6.4 不失为经典极值条件的完美推广.但从应用上考虑,它未必能令人满意.首先,要求 x_0 为内点就排除了 x_0 为边界点或 D 根本无内点的情况,而这在应用上恰恰是常见的情况.其次,如我们在 §4.1 中已指出的,在无限维空间中要求函数可微是一个很强的限制,更不必说要求 f 是 C^2 函数了.这就促使我们去寻求较弱的微分条件,基本想法是将 F 导数代之以方向导数甚至次微分.

定理 4.6.5 设 $D \subset X$ 是凸集, $x_0 \in D, f: D \to \mathbf{R}$.

(i) 必要条件:若 x_0 是 f 在 D 上的局部极小点,对任给 $h \in D - x_0$,方向导数 $\mathrm{D}_+ f(x_0, h)$ (依式(4.1.18))恒存在,则 $\mathrm{D}_+ f(x_0, D - x_0) \geqslant 0$,即

$$\mathrm{D}_+ f(x_0, x - x_0) \geqslant 0 \quad (\forall\, x \in D). \tag{4.6.7}$$

(ii) 充分条件:若 f 是凸函数,且条件(4.6.7)满足,则 x_0 是 f 在 D 上的最小值点.

证 (i) 取定 $x \in D$,令 $x_t = t' x_0 + tx, t \in (0,1)$,则

$$\begin{aligned} 0 &\leqslant \lim_{t \downarrow 0} \frac{f(x_t) - f(x_0)}{t} \qquad (f(x_0)\text{ 为局部极小}) \\ &= \lim_{t \downarrow 0} \frac{f(x_0 + t(x - x_0)) - f(x_0)}{t} \\ &= \mathrm{D}_+ f(x_0, x - x_0). \end{aligned}$$

这表明条件(4.6.7)满足.

(ii) 任取 $x \in D$,有

$$\begin{aligned} 0 &\leqslant \mathrm{D}_+ f(x_0, x - x_0) \quad (\text{用式(4.6.7)}) \\ &\leqslant \Delta f(x_0, x - x_0) \quad (\text{用式(4.5.5)}) \\ &= f(x) - f(x_0), \end{aligned}$$

这正表明 $f(x_0) \leqslant f(x) (\forall\, x \in D)$,即 $f(x_0)$ 是最小值. □

若 x_0 是 D 的内点,且 $f'(x_0)$ 或 $\mathrm{D}f(x)$ 存在,则条件(4.6.7)等价于 $f'(x_0) = 0$ 或 $\mathrm{D}f(x_0) = 0$(请试证).在这种情况下,定理 4.6.5(i)重合于 4.6.4(i),而定理 4.6.5(ii)表明:对于凸函数 $f(x)$, $f'(x_0) = \theta$ (或 $\mathrm{D}f(x_0) = 0$)$\Leftrightarrow f(x_0)$ 是最小值.注意 $f'(x_0) = 0$ 相当于 f 的图形在 x_0 有水平切平面,因此以上结论在直观上是很自然的.若 D 是 X 中的线性流形(这意味着 $D - x_0$ 是 X 的子空间),

$Df(x_0,h)$ 恒存在,则条件(4.6.7)相当于 $Df(x,h)=0(\forall h\in D-x_0)$. 若进而设 G 导数 $Df(x_0)$ 存在,则条件(4.6.7)意味着 $Df(x_0)\in(D-x_0)^{\perp}$. 这些事实在定理 4.6.5 的应用中是值得注意的.

当然,定理 4.6.5 的价值应体现于 x_0 非内点或 f 在 x_0 不可微的情况. 最简单的例子是定义于区间 $[x_0,b)$ 上的实函数 $f(x)$. $\forall x\in(x_0,b)$, 有

$$D_+ f(x_0,x-x_0)=\lim_{t\downarrow 0}\frac{f(x_0+t(x-x_0))-f(x_0)}{t}$$
$$=f'_+(x_0)(x-x_0).$$

因此条件(4.6.7)相当于 $f'_+(x_0)\geqslant 0$. 这就由定理 4.6.5 得出:若 f 在定义区间的左端点 x_0 取得局部极小值,则 $f'_+(x_0)\geqslant 0$; 类似地,若 $f(x)$ 在定义区间的右端点 x_0 取得局部极小值,则 $f'_-(x_0)\leqslant 0$, 假定如上的单侧导数存在.

下面是应用定理 4.6.5 的更有趣的例子.

例 4.6.6(变分问题) 考虑如下的最小化问题:

$$\begin{cases}\min\int_a^b F(t,x(t),x'(t))\mathrm{d}t,\\ \text{s.t. } x(a)=\alpha,x(b)=\beta,x\in X=C^2(J),J=[a,b].\end{cases}\tag{4.6.8}$$

令 $D=\{x\in X: x(a)=\alpha,x(b)=\beta\}$, 定义

$$f(x)=\int_a^b F(t,x(t),x'(t))\mathrm{d}t\quad(x\in D),$$

则问题(4.6.8)相当于求泛函 f 在 D 上的极小点. 假定 $F\in C^2$, 并以 F_1,F_2,F_3 记 F 的一阶偏导数, F_{11} 等记二阶偏导数. 今用定理 4.6.5 来表出问题(4.6.8)的解应满足的条件. 首先求出方向导数,为此用如下较简易的方法(更严格的方法亦不困难):任给 $h\in X$, 有(参考例 4.1.4)

$$\begin{aligned}Df(x,h)&=\frac{\mathrm{d}}{\mathrm{d}s}f(x+sh)\Big|_{s=0}\\&=\int_a^b\left[\frac{\mathrm{d}}{\mathrm{d}s}F(t,x(t)+sh(t),x'(t)+sh'(t))\Big|_{s=0}\right]\mathrm{d}t\\&=\int_a^b[F_2(t,x(t),x'(t))h(t)+F_3(t,x(t),x'(t))h'(t)]\mathrm{d}t.\end{aligned}$$

D 显然是 X 中的线性流形,它是子空间 $D_0=\{x\in X: x(a)=x(b)=0\}$ 平移的结果. 依定理 4.6.5(i)(参考其后的说明), f 在 x 取得极小的必要条件是

$Df(x,h)=0(\forall h\in D_0)$，即 $\forall h\in D_0$，有

$$0=\int_a^b F_2(t,x(t),x'(t))h(t)dt+\int_a^b F_3(t,x(t),x'(t))h'(t)dt=\int_a^b\left[F_2(t,x(t),x'(t))-\frac{d}{dt}F_3(t,x(t),x'(t))\right]h(t)dt,$$

其中用了分部积分. 上述积分的右端可写成

$$\langle\varphi,h\rangle=\int_a^b\varphi(t)h(t)dt,$$

其中 $\varphi=F_2(t,x,x')-\frac{d}{dt}F_3(t,x,x')\in C(J)$. 因 D_0 在 $L^2(J)$ 中稠密(参考1.3.8)，故由 $\varphi\in D_0^\perp$ 得出 $\varphi=0$，这意味着

$$F_{33}x''+F_{23}x'+F_{13}-F_2=0, \tag{4.6.9}$$

其中 $F_2=F_2(t,x(t),x'(t))$，F_{13}，F_{23}，F_{33} 仿此. 方程(4.6.9)称为变分问题(4.6.8)的 **Euler 方程**[①]. 当 F 不显含 t 时，方程(4.6.9)可简化为

$$\frac{d}{dt}(F-F_3x')=0. \tag{4.6.9$'$}$$

以上结果在变分学中是熟知的且被广泛应用.

今用以上结论来解著名的**最速降线问题**. 设 y 轴铅直向下. 求连结原点(0,0)到点 $(b,h)(b,h>0)$ 的光滑曲线 $y=y(x)$，使沿该曲线无摩擦地下滑的质点所用时间 T 最小. 不妨设质点的质量为 1. 因质点动能的增加等于其势能的减少，故成立

$$\frac{1}{2}v^2=gy,$$

其中 v 是速度，而 g 为重力加速度. 因

$$T=\int_0^T dt=\int_0^T\frac{ds(t)}{v}=\int_0^b\frac{\sqrt{1+y'^2}}{2gy}dx,$$

故问题归于解最小化问题：

---

① 泛函极值问题的必要条件通常表为一个(或一组)方程，统称为 Euler 方程. 例如，本节的方程(4.6.12)与(4.6.15)都可称为 Euler 方程. Euler 方程经常取微分方程的形式.

$$\begin{cases}\min\int_0^b F(y,y')\mathrm{d}x \\ \text{s.t. } y(0)=0, \quad y(b)=h,\end{cases}$$

其中 $F(y,y')=\sqrt{(1+y'^2)/y}$. 分别 x,y,y' 代 t,x,x', 从方程(4.6.9)′得

$$y(1+y'^2)=2a,$$

a 是正常数. 令 $y'=\cot(t/2)$, 则

$$y=\frac{2a}{1+y'^2}=a(1-\cos t);$$

$$\begin{aligned}x&=\int_0^t \tan\frac{t}{2}\mathrm{d}y(t)=a\int_0^t(1-\cos t)\mathrm{d}t\\&=a(t-\sin t).\end{aligned}$$

可见,所求曲线是熟知的旋轮线(又称摆线).

定理 4.6.5 已经大大降低了可微性条件. 再往前走一步,就要用到次微分了.

定理 4.6.7 设 $D\subset X, f: X\to\mathbf{R}, x_0\in D$.

(i) 充要条件: x_0 是 f 在 D 上的最小值点 $\Leftrightarrow 0\in\partial(f\mid D)(x_0)$.

(ii) 充分条件:若 $\exists\, u\in\partial f(x_0)$, 使得 $\langle u, D-x_0\rangle\geqslant 0$, 即

$$\langle u, x-x_0\rangle\geqslant 0 \quad (\forall x\in D), \tag{4.6.10}$$

则 x_0 是 f 在 D 上的最小值点.

从形式上看,条件 $0\in\partial(f\mid D)(x_0)$ 对应于定理 4.6.4 中的条件 $f'(x_0)=0$, 而条件(4.6.10) 则对应于定理 4.6.5 中的条件(4.6.7).

证 (i) 直接看出:

$$\begin{aligned}f(x_0)\leqslant f(x)(\forall x\in D)&\Leftrightarrow 0\leqslant f(x)-f(x_0) \quad (\forall x\in D)\\&\Leftrightarrow 0\in\partial(f\mid D)(x_0).\end{aligned}$$

(ii) 若 $u\in\partial f(x_0)$ 满足条件(4.6.10),则

$$0\leqslant\langle u, x-x_0\rangle\leqslant f(x)-f(x_0), \quad (\forall x\in D)$$

这表明 $f(x_0)$ 是最小值. □

§4.6.3 条件极值

最小化问题(4.6.3)中的集 D, 通常由一定的关系式界定. 这类关系式既可以

是等式，也可以是不等式，统称为极值问题的约束条件. 下面仅考虑带等式约束的最小化问题，这类问题正是微分学中所处理的条件极值问题的自然推广，其一般形式可表为

$$\min_{x \in X} f(x), \quad \text{s.t.}\ g(x) = 0, \tag{4.6.11}$$

其中 $f: X \to \mathbf{R}$ 称为**目标函数**，$g: X \to Y$ 称为**约束函数**[①]，$g(x) = 0$ 就是**约束条件**. 若令 $D = \{x \in X : g(x) = 0\}$，则问题(4.6.11)可改写为(4.6.3)的形式. 因此，定义 4.6.1 中所给出的概念与术语全部适用于问题(4.6.11).

关于条件极值的基本课题之一，是给出类似于定理 4.6.4 的微分条件. 下面是一个典型结果.

定理 4.6.8　设 x_0 是问题(4.6.11)的局部最优解，f 在 x_0 可微，g 在 x_0 邻近为 C^1 映射，$g'(x_0)$ 是满射. 则存在 $\lambda \in Y^*$，使得

$$f'(x_0) + \lambda \circ g'(x_0) = 0. \tag{4.6.12}$$

证　令 $T = g'(x_0)$，则式(4.6.12)相当于

$$f'(x_0) = -T^* \lambda \in R(T^*).$$

由定理条件有 $R(T) = Y$，这推出 $R(T^*) = N(T)^\perp$ (用定理 2.4.13). 于是只要证 $f'(x_0) \in N(T)^\perp$. 取定 $h \in N(T)$，今证 $\langle f'(x_0), h \rangle = 0$. 由下文中补述的 Lusternik 定理，当 $t \in \mathbf{R}$，$|t|$ 充分小时，方程

$$g(x_0 + y + th) = 0$$

有解 $y = y(t)$，且当 $t \to 0$ 时 $y = o(t)$. 于是当 $|t|$ 充分小时，

$$\begin{aligned} 0 &\leqslant f(x_0 + y + th) - f(x) && (x_0 \text{ 为极小点}) \\ &= \langle f'(x_0), y + th \rangle + o(y + th) \\ &= t\langle f'(x_0), h \rangle + o(t). && (\text{用 } y = o(t)) \end{aligned}$$

因上式右端的符号决定于第一项，而 t 可正可负，故必 $\langle f'(x_0), h \rangle = 0$，如所要证.

□

上述证明中用到如下定理.

Lusternik 定理　设 $g: X \to Y$ 是 C^1 映射，$x_0 \in X$，$g'(x_0)$ 是满射. 则当 $z \in X$，$\|z\|$ 充分小时，方程

---

① 只需假定 f, g 定义于某个开集 $\Omega \subset X$ 内就行了. 此处取 $\Omega = X$，只是为了简便.

$$g(x_0+y+z)=g(x_0)+g'(x_0)z$$

有解 $y=y(z)$，且当 $z\to 0$ 时，$y=o(z)$.

现在来解释一下定理 4.6.8 的意义. 令

$$L(x,\lambda)=f(x)+\langle\lambda,g(x)\rangle\quad(x\in X,\lambda\in Y^*),\tag{4.6.13}$$

则式(4.6.12)可写作

$$L_x(x_0,\lambda)=0.\tag{4.6.12$'$}$$

从形式上看，式(4.6.12)′似乎是无条件极值问题

$$\min L(x,\lambda),\quad x\in X\tag{4.6.14}$$

的必要条件(参看定理 4.6.4). 函数(4.6.13)称为问题(4.6.11)的 **Lagrange 函数**，而使式(4.6.12)′成立的 $\lambda\in Y^*$ 称为 **Lagrange 乘子**. 这样，通过引进 Lagrange 函数，形式上将一个条件极值问题转化成了一个无条件极值问题；原问题中的约束条件已被吸收到 Lagrange 函数 $L(x,\lambda)$ 中了. 极值问题中约束条件的形式可能有许多变化，但构作 Lagrange 函数的思想则具有极大的普遍性.

若在问题(4.6.11)中取 $Y=\mathbf{R}^m$，$g=(g_1,g_2,\cdots,g_m)^{\mathrm{T}}$，并令 $\lambda=(\lambda_i)\in(\mathbf{R}^m)^*=\mathbf{R}^m$，则 Lagrange 函数(4.6.13)相应地写成

$$L(x,\lambda)=f(x)+\sum_i\lambda_i g_i(x).$$

这正是微分学中所熟知的形式，其中 λ_i 就是通常所说的 Lagrange 乘数. 因

$$g'(x_0)=[g'_1(x_0),g'_2(x_0),\cdots,g'_m(x_0)]^{\mathrm{T}},$$

故条件(4.6.12)可写成

$$f'(x_0)+\sum_i\lambda_i g'_i(x_0)=0,\tag{4.6.15}$$

而条件 $R(g'(x_0))=\mathbf{R}^m$ 相当于 $\{g'_i(x_0):1\leqslant i\leqslant m\}$ 线性无关. 若 $X=\mathbf{R}^n$，则 $g'(x_0)\in\mathbf{R}^{m\times n}$，于是 $g'(x_0)$ 为满射 $\Leftrightarrow$ rank $g'(x_0)=m$.

应用以上结论的一个简单而有趣的例子是例 1.5.11. 若令

$$f(x)=\|x\|,g_i(x)=\langle x,a_i\rangle-\alpha_i\quad(1\leqslant i\leqslant n),$$

则可将例 1.5.11 中的问题(1.5.32)写成现在考虑的标准形式

$$\min_{x\in H}f(x),\quad \text{s.t.}\ g_i(x)=0\quad(1\leqslant i\leqslant n).$$

因 $f'(x)=x/\|x\|$，$g'_i(x)=a_i$，故条件(4.6.15)可写成(以 x 代 x_0)

$$\frac{x}{\|x\|}+\sum_i\lambda_i a_i=0.$$

令 $\beta_i=-\lambda_i\|x\|$，则得到 $x=\sum\beta_i a_i$. 以此代入方程组

$$\langle x,a_i\rangle=\alpha_i\quad(1\leqslant i\leqslant n),$$

即得到关于 β_i 的线性方程组，所得结果与例 1.5.11 恰好一致.

关于条件极值的一个更具体的例子是**等周问题**：在 xy 平面上求一光滑闭曲线

$$x=x(t),\quad y=y(t)\quad(0\leqslant t\leqslant 2\pi),$$

使其周长一定，而其所围面积达到最大. 这意味着解以下条件最大化问题：

$$\begin{cases}\max\displaystyle\int_0^{2\pi}[x(t)y'(t)-x'(t)y(t)]\mathrm{d}t,\\ \text{s.t.}\displaystyle\int_0^{2\pi}\sqrt{x'^2(t)+y'^2(t)}\mathrm{d}t=\text{const},\end{cases}$$

其中 $x(t),y(t)$ 是以 2π 为周期的 C^2 函数. 注意，我们用了由 Green 公式推出的面积公式

$$L\text{ 所围面积}=\frac{1}{2}\int_L x\mathrm{d}y-y\mathrm{d}x.$$

作 Lagrange 函数：

$$L(x,y,\lambda)=\int_0^{2\pi}[x(t)y'(t)-x'(t)y(t)+\lambda\sqrt{x'^2(t)+y'^2(t)}]\mathrm{d}t.$$

用类似于例 4.1.4 中所用的方法可求得

$$L_x(x,y,\lambda)h=\int_0^{2\pi}[2-\lambda k(t)]y'(t)h(t)\mathrm{d}t,$$

$$L_y(x,y,\lambda)h=\int_0^{2\pi}[\lambda k(t)-2]x'(t)h(t)\mathrm{d}t,$$

其中

$$k=(x''y'-x'y'')(x'^2+y'^2)^{-3/2},$$

$|k|$ 正是曲线的曲率. 于是，由 Euler 方程

$$L_x(x,y,\lambda)=0,\quad L_y(x,y,\lambda)=0$$

得出 $k=2/\lambda=\text{const}$. 可见，所求的曲线就是圆周，从直观上看这是意料中的事.

§4.6.4 最佳逼近

本书中已有几处涉及最佳逼近问题(参看推论1.4.5与定理1.5.6),现在我们从极值理论的角度作一新的考虑.

给定 $a\in X$ 与非空集 $D\subset X$,设 $a\bar{\in} D$. a 在 D 中的最佳逼近,可定义为最小化问题

$$\min\|x-a\|,\quad x\in D \tag{4.6.16}$$

的最优解. 问题(4.6.16)的目标函数 $f(x)=\|x-a\|$ 似乎是平凡的. 但因集 D 有各种选择的可能性,问题(4.6.16)未必是平凡的. 如你已看到的,即使对于 D 为子空间这种极规范的情况,最佳逼近问题的求解通常亦不容易. 深入的讨论远不是此处所宜进行的,下面只是将本节已建立的那些命题用到问题(4.6.16),看看能得到哪些特殊的结论. 至少,这能为本节所建立的极值条件的应用提供简单的例证.

(i) 函数 $f(x)=\|x-a\|$ 是连续凸函数(参看例4.5.2(i)),且显然满足条件(4.6.4). 因此由定理4.6.2得出:若 D 是自反空间 X 中的闭凸集,或 D 是 X 的某个有限维子空间中的闭集,则 a 在 D 中的最佳逼近存在;当 D 为凸集时,a 在 D 中的最佳逼近之全体构成一凸集. 以上结论可与推论1.4.5及定理1.5.6(i)相对照. 对于应用上常见的情况,取 D 为 X 的有限维子空间,因而最佳逼近问题必有解,且解之全体是一凸集.

(ii) 设 X 是实 Hilbert 空间,D 是凸集,由式(4.1.3),$\forall x\in D$,有

$$\nabla\|x-a\|=(x-a)/\|x-a\|.$$

于是由定理4.6.5得出:$x_0\in D$ 是 a 的最佳逼近的充要条件是

$$\langle\|x_0-a\|^{-1}(x_0-a),x-x_0\rangle\geqslant 0,$$

即

$$\langle a-x_0,x-x_0\rangle\leqslant 0\quad(\forall x\in D). \tag{4.6.17}$$

几何上,条件(4.6.17)意味着:$\forall x\in D$,向量 $x-x_0$ 与 $a-x_0$ 张成钝角(看图4-3). 这无疑正好合于我们从平常空间中获得的直觉经验. 若 D 是 X 的子空间,则条件(4.6.17)相当于

$$\langle a-x_0,x\rangle\leqslant 0\quad(\forall x\in D).$$

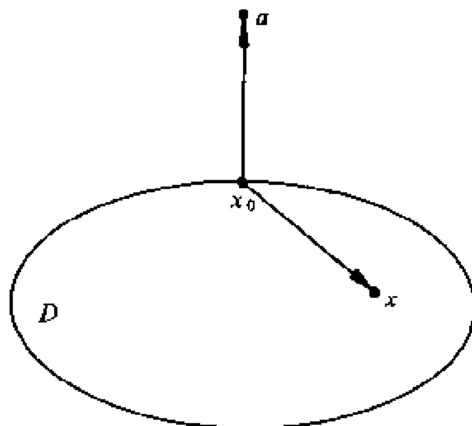


图4-3

因 x 可代以 $-x$，故以上条件又可代以 $\langle a-x_0,x\rangle\geqslant 0(\forall x\in D)$. 这就表明式(4.6.17)等价于 $x_0-a\in D^{\perp}$，这与定理 1.5.6 的结论一致.

(iii) 对任何 $x\in D$，由式(4.5.11)有

$$\partial\|x-a\|=\{u\in X^*: u(x-a)=\|x-a\|,\|u\|=1\}.$$

于是由定理 4.6.7 推出：若 $x_0\in D$，存在 $u\in X^*$，使得

$$\begin{cases}u(x_0-a)=\|x_0-a\|, \quad \|u\|=1,\\ u(x-x_0)\geqslant 0 \quad (\forall x\in D),\end{cases} \tag{4.6.18}$$

则 x_0 是 a 在 D 中的最佳逼近. 对于一般的集 D，未必容易找到满足条件(4.6.18)的 $u\in X^*$，但上述条件毕竟为解最佳逼近问题开辟了一条途径. 若 D 是 X 的子空间，则式(4.6.18)中的不等式显然等价于 $u(x)=0(\forall x\in D)$，即 $u\in D^{\perp}$. 于是条件(4.6.18)可改写成

$$u\in D^{\perp}, \quad \|u\|=1, \quad u(x_0-a)=\|x_0-a\|. \tag{4.6.18$'$}$$

其中的等式 $u(x_0-a)=\|x_0-a\|$，有些作者称为共线条件：u 与 x_0-a 共线. 这一说法在 Hilbert 空间中更加贴切. 若 X 是 Hilbert 空间，x_0 是 a 在 D 上的正投影，则取

$$u=(x_0-a)/\|x_0-a\|$$

恰使式(4.6.18)$'$满足，因此 x_0 是 a 在 D 中的最佳逼近. 显然，此时 u 与 x_0-a 确实共线.

(iv) 设 X 是 Hilbert 空间，D 是 X 的闭子空间，P 是从 X 到 $D^{\perp}$ 的正投影，则可将问题(4.6.16)表为条件极值问题：

$$\min\|x-a\|, \quad \text{s.t. } Px=0. \tag{4.6.19}$$

注意 $P: X\to D^{\perp}$ 是有界线性算子且为满射. 于是由定理 4.6.8 推出：若 x_0 是 a 在 D 中的最佳逼近，则存在 $\lambda\in(D^{\perp})^*=D^{\perp}$，使得

$$\frac{x_0-a}{\|x_0-a\|}+P\lambda=0,$$

即

$$\frac{x_0-a}{\|x_0-a\|}=-P\lambda=-\lambda\in D^{\perp},$$

从而 $x_0-a\in D^{\perp}$. 这再次得出定理 1.5.6 中的结论.

习 题

1. 设 $F: X \to Y$ 可微，$F(\alpha x) = \alpha Fx(x \in X, \alpha \in \mathbf{R})$，证明 $F \in L(X,Y)$.

2. 设 $f: [a,b] \times \Omega \subset \mathbf{R} \times X \to Y$ 连续且 $f_x(t,x)$ 连续，$Fx = \int_a^b f(t,x)\mathrm{d}t$，求 $F'(x)$.

3. 设 $J = [a,b], f: J \times \mathbf{R} \to \mathbf{R}$ 连续且 $f_x(t,x)$ 连续，$Fu = \int_a^b f(t,u(t))\mathrm{d}t, u \in C(J)$，求 $F'(u)$.

4. 设 $J = [0,1], f(u) = \int_0^1 |u(x)| \mathrm{d}x (u \in C(J))$，证明 $f'(0)$ 不存在.

5. 设 $T: X \times Y \to Z$ 是双线性的且 $\| T(x,y) \| \leqslant k \| x \| \| y \|$，求 $T'(x,y), T''(x,y)$；证明 $T \in C^{\infty}$.

6. 设 $H = G \circ F, G$ 与 F 二次可微，求 $H''(x)hk$.

7. 设 $F(x,y): \Omega \subset X \times Y \to Z$ 是对 x 连续，对 y 可微，$\sup \| F_y(x,y) \| < \infty$，证明 $F \in C(\Omega, Z)$.

8. 设 $F: \Omega \subset X \to Y$ 可微，Ω 是凸开集，证明 $\mathrm{Lip}F = \sup_{x \in \Omega} \| F'(x) \|$.

9. 设 $f: \mathbf{R}^n \to X$ 是 C^r 函数，证明 r 阶偏导数 $\partial^r f / \partial x_1^{i_1} \cdots \partial x_n^{i_n} (i_1 + i_2 + \cdots + i_n = r)$ 与求导顺序无关.

10. 设 $f: \Omega \subset X \to \mathbf{R}$ 在 Ω 内 $n+1$ 次可微，线段 $[x, x+h] \subset \Omega$，证明

$$f(x+h) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(x)h^k + \frac{1}{(n+1)!} f^{(n+1)}(x + \theta h)h^{n+1} (0 < \theta < 1).$$

11. 设 $F: B_2(0) \subset X \to Y$ 是 C^2 映射，$\| Fx \| \leqslant 1, \| F''(x) \| \leqslant 1$. 证明当 $\| x \| \leqslant 1$ 时 $\| f'(x) \| \leqslant 5/2$.

12. 设 $F, F_n: D \subset X \to Y, F_n x \to Fx (n \to \infty, x \in D)$，证明 $\mathrm{Lip}F \leqslant \varliminf_n \mathrm{Lip}F_n$.

13. 设 $F: B_\rho(x_0) \subset X \to X, \mathrm{Lip}F \leqslant r < 1, \| x_0 - Fx_0 \| \leqslant \rho(1-r)$，证明 $F^n x_0 \to x^* = Fx^*$.

14. 设 $D \subset X$ 是非空闭集，$F: D \to D, \sum \mathrm{Lip}F^n < \infty, x_0 \in D$，则 $F^n x_0 \to x^* = Fx^*$，估计 $\| F^n x_0 - x^* \|$.

15. 设 $D \subset X$ 是非空紧集，$F: D \to D$ 满足 $\| Fx - Fy \| < \| x - y \| (x \neq y)$，证明 F 有惟一不动点.

16. 验证 $Fx = x + x^{-1}$ 在 $[1, \infty)$ 上满足 $| Fx - Fy | < | x - y | (x \neq y)$，但无不动点.

17. 设 $K(\cdot,\cdot) \in C(J \times J), g \in C(\mathbf{R}), v \in C(J), J = [a,b]$，

$$\max_{x \in J} \int_a^b | K(x,y) | \mathrm{d}y < 1/\mathrm{Lip}g,$$

证明方程 $u(x) = \int_a^b K(x,y)g(u(y))\mathrm{d}y + v(x)$ 有惟一解 $u \in C(J)$.

18. 设 $J: G \triangleq \mathrm{GL}(X) \to L(X), A \to A^{-1}$，证明

$$J^{(n)}(A)B^n = (-1)^n n!A^{-1}(BA^{-1})^n, A \in G, B \in L(X).$$

19. 设 $U \subset X$ 与 $V \subset Y$ 是开集，$F : U \to V$ 与 $F^{-1} : V \to U$ 均为 C^1 映射，证明：$F'(x) : X \cong Y (\forall x \in U)$.

20. 设 $F : \Omega \subset X \times Y \to Z$ 是 C^r 映射，$r \geqslant 1, (x_0, y_0) \in \Omega, F_y(x_0, y_0) : Y \cong Z, z_0 = F(x_0, y_0)$. 证明存在 (x_0, z_0) 邻近的 C^r 映射 $f(x, z)$，使得 $F(x, f(x, z)) = z, f(x_0, y_0) = y_0$.

21. 设 $F : X \to Y$ 是紧算子，$T \in L(X, Y)$，当 $\| x \| \to \infty$ 时 $\| Fx - Tx \| = o(\| x \|)$. 证明 $T \in \mathrm{CL}(X, Y)$.

22. 设 $D \subset X$ 是有界闭集，$F : D \to X$ 是紧算子，证明 $(I - F)D$ 是闭集.

23. 设 $D \subset X$ 是紧凸集，$F \in C(D, D)$，证明 F 有不动点.

24. 设 $D \subset X$ 是有界闭集，$F : D \to X$ 是紧算子，$\lambda_n \in (0,1), x_n \in D, \lambda_n x_n = Fx_n$，$\inf\limits_n \| Fx_n \| > 0$. 证明存在 $\lambda \in (0,1]$，使得方程 $\lambda x = Fx$ 有解.

25. 设 $f : D \subset X \to \mathbf{R}$ 是凸函数，证明 $\varphi(t) \triangleq f(x + th)$ 在其有定义的区间上为凸函数.

26. 设 $f : \Omega \subset X \to \mathbf{R}$ 是 C^2 的凸函数，证明 $f''(x)h^2 \geqslant 0 (x \in \Omega, h \in X)$.

27. 设 $f : X \to \mathbf{R}$ 连续，$f\left(\frac{x + y}{2}\right) \leqslant \frac{1}{2}[f(x) + f(y)]$，证明 f 为凸函数.

28. 求 $\partial \| x \|^2 (x \in X)$.

29. 设 $u \in \partial f(x), v \in \partial f(y)$，验证 $\langle u - v, x - y \rangle \geqslant 0$.

30. 设 f 为凸函数，$\mathrm{D}f(x, h)$ 恒存在，证明 $\mathrm{D}f(x, h)$ 对 h 为凸函数.

习题答案与提示

第一章

1. 由 $\|x-a\| \leqslant \|x-y\| + \|y-a\|$ $(\forall a \in A)$ 推出 $d(x,A) \leqslant \|x-y\| + d(y,A)$；同理 $d(y,A) \leqslant \|x-y\| + d(x,A)$，二者合起来得所要证.

4. 若 $x^{(n)} = (x_1^{(n)}, x_2^{(n)}) \in X_1 \times X_2$ 依积范数是 Cauchy 列，则 $\{x_i^{(n)}\}$ 是 $X_i (i=1,2)$ 中的 Cauchy 列，于是由 $x_i^{(n)} \to x_i (i=1,2)$ 推出 $x^{(n)} \to (x_1, x_2)(n \to \infty)$.

5. 若 $\alpha\|x\| \leqslant \|Tx\| \leqslant \beta\|x\| (\forall x \in X), \alpha, \beta > 0$，则 $x_n \to x \Leftrightarrow \|x_n - x\| \to 0 \Leftrightarrow \|T(x_n - x)\| \to 0(\|T(x_n - x)\| \leqslant \beta\|x_n - x\|, \|x_n - x\| \leqslant \alpha^{-1}\|T(x_n - x)\|) \Leftrightarrow Tx_n \to Tx(n \to \infty)$. 反之，若 T 非拓扑同构，不妨设 $\|Tx\| \leqslant \mathrm{const}\|x\|$ 不成立，则有 $\{x_n\} \subset X$，使 $\|Tx_n\| > n\|x_n\| (\forall n \geqslant 1)$. 令 $y_n = x_n / n\|x_n\|$，则 $y_n \to 0$，而 Ty_n 不收敛于 $T0$.

6. 以 A 表 $[0,1]$ 上次数不大于 n 的多项式之全体，则 $T: A \to \mathbf{K}^{n+1}, \sum_0^n \alpha_i x^i \to (\alpha_i)$ 是线性同构，因而可用定理 1.1.7 与式(1.1.12).

7. 以 A 表 $[0,1]$ 上次数不大于 5 的多项式之全体，则 A 是 $C[0,1]$ 的闭子空间.

8. 令 $A = \mathrm{span}\{x_1, x_2, \cdots, x_n\}$，则 $T: \mathbf{K}^n \to A, (\lambda_i) \to \sum \lambda_i x_i$ 是线性同构，因而可用定理 1.1.7 与式(1.1.12)(对照题 6).

9. 将 $(x_1, \cdots, x_n)^{\mathrm{T}} \in \mathbf{K}^n$ 等同于 $(x_1, \cdots, x_n, 0, \cdots)^{\mathrm{T}} \in l^p$，可以认为 $\mathbf{K}^n$ 是 l^p 的子空间，且是闭子空间，因而是 Banach 空间. $\mathbf{K}^n$ 中的范数收敛(无论什么范数)总归结为依坐标收敛.

10. 只要说明 $C_0(\mathbf{R})$ 是 $B(\mathbf{R})$ 的闭子空间. 设 $\{u_n\} \subset C_0(\mathbf{R}), u_n \rightrightarrows u$，则必 $u \in C(\mathbf{R})$，由 $|u(x)| \leqslant |u_n(x)| + \|u_n - u\|_0$ 看出 $\lim\limits_{|x| \to \infty} u(x) = 0$.

11. 是赋范空间而非 Banach 空间，因一致收敛的多项式序列的极限不必为多项式.

12. 设 $J = [a,b]$，则 $\|u_n - u\|_p^p \leqslant (b-a)\|u_n - u\|_0^p$.

13. 可以断定 $u(x)$ 在 $[0,1]$ 上无零点：取 X 为$[0,1]$上次数不大于 n 的多项式之全体，$A = \{u \in X : u$ 在$[0,1]$上无零点$\}$，则 A 是开集，因而微小扰动不改变 $u \in A$ 这一关系.

14. 否则 A 包含一个球 $B_r(a), r > 0$，这推出 $B_r(0) = B_r(a) - a \subset A$，从而 $X = \bigcup_1^\infty nB_r(0) \subset A$.

15. 不能保证 $[0,1]$ 上的 2 次多项式之全体 A 在空间 $C[0,1]$ 中无内点(参考题 14)，关系 $u \in A$ 未必能承受微小扰动.

16. $d(\cdot, A)$ 是 X 上的连续函数(参考题 1).

17. $\mathbf{R}^2$ 中每个正方形可等分为边长充分小的正方形；边长为 a 的正方形含于半径为 $a/\sqrt{2}$ 的圆.

18. 取 $a_n \in A, b_n \in B$，使 $\|a_n - b_n\| \to d(A,B)$，不妨设 $\{a_n\}$ 与 $\{b_n\}$ 均收敛.

19. 否则，$\forall n\in\mathbf{N}$，$\exists x_n$，使 $d(x_n,A)<1/2n$，$d(x_n,B)<1/2n$；于是有 $a_n\in A$，$b_n\in B$，使 $\|a_n-b_n\|\leqslant\|x_n-a_n\|+\|x_n-b_n\|<1/n$. 不妨设 $a_n\to a\in A$，则必 $b_n\to a\in B$.

20. 取 $y\in X\setminus A$，$a\in A$，使 $\|y-a\|=d(y,A)$（用推论 1.4.5），则 $x=(y-a)/\|y-a\|$ 为所求.

21. 由 $|f(u)-f(v)|\leqslant\|u-v\|_1\leqslant\|u-v\|_0$ 知 $f(\cdot)$ 连续. $f(1)=1$ 是 f 在 S 上的最大值，但 f 在 S 上达不到最小值；设 $u_n(x)$ 在 $[0,1/n]$ 上是线性函数，在 $[1/n,1]$ 上为零，$u_n(0)=1$，则 $u_n\in S$，而

$$f(u_n)=\int_0^{1/n}u_n(x)\mathrm{d}x=\frac{1}{2n}\to 0.$$

但 f 在 S 上不取零值.

22. 由 $|\tilde{u}(x)-\tilde{u}(y)|\leqslant\text{const}\,|x-y|$ 说明 $\{\tilde{u}:u\in A\}$ 等度连续.

23. 令 $A_n=\text{span}\{x_1,x_2,\cdots,x_n\}$，则 $X=\bigcup_1^{\infty}A_n$，A_n 是闭疏集（用题 14），因而 X 是第一纲的.

24. 指明该集内部非空，然后用定理 1.4.11.

25. 指明该集是闭集且无内点. 例如，设 $J=[0,1]$，$0\leqslant u\in L^p(J)$，令 $u_n=u\chi_{[1/n,1]}-\chi_{[0,1/n]}$，则 $u_n\xrightarrow{L^p}u$，但 u_n 不是非负的.

26. 显然 $x_n\to x\Rightarrow\lim\limits_{m,n}\langle x_m,x_n\rangle=\|x\|^2$. 反之，若 $\lim\limits_{m,n}\langle x_m,x_n\rangle=\alpha$，则 $\|x_m-x_n\|^2\to 0\,(m,n\to\infty)$.

27. 指明 $\{x_i\}$ 线性相关 $\Leftrightarrow$Gram 矩阵之各行线性相关.

28. 取 $u(x)=x$，$v(x)=1-x$ 说明中线公式不成立.

29. 展开 $\|x+\alpha y\|^2=\|x-\alpha y\|^2$ 并以 $\alpha=\langle x,y\rangle$ 代入. 若 $\mathbf{K}=\mathbf{R}$，则 $x\perp y\Leftrightarrow\|x+y\|=\|x-y\|$.

30. 展开 $\|x+\alpha y\|^2\geqslant\|x\|^2$ 并以 $\alpha=-\langle x,y\rangle/\|y\|^2$ 代入.

31. $L_0=1$，$L_1=x\sqrt{3/2}$，$L_2=\sqrt{5/8}(3x^2-1)$.

32. 将 $\{1,x,x^2\}$ 关于 $e^{-x^2}\mathrm{d}x$ 在 $(-\infty,\infty)$ 上标准正交化为 $\{e_0,e_1,e_2\}$，然后用 $H_n=(2^n n!\sqrt{\pi})^{1/2}e_n$ 求得 $H_0=1$，$H_1=2x$，$H_2=4x^2-2$.

33. $\dfrac{6\pi-15}{4}-\dfrac{15(\pi-3)}{4}x^2\approx 0.9624-0.5310x^2$.

34. $\dfrac{2}{\pi}-\dfrac{5}{\pi^5}(24-2\pi^2)(6x^2-6\pi x+\pi^2)\approx -0.4177x^2+1.3123x-0.0505$.

35. 令 $\varphi_1=\chi_{(0,1/3)}$，$\varphi_2=\chi_{(1/3,2/3)}$，$\varphi_3=\chi_{(2/3,1)}$，则 $\{\sqrt{3}\varphi_1,\sqrt{3}\varphi_2,\sqrt{3}\varphi_3\}$ 是标准正交系，所求最佳逼近为 $3[(e^{1/3}-1)\varphi_1+(e^{2/3}-e^{1/3})\varphi_2+(e-e^{2/3})\varphi_3]\approx 1.1868\varphi_1+1.6564\varphi_2+2.3117\varphi_3$.

第二章

1. $TB_1(0)\subset\overline{B}_\beta(0)\Rightarrow\forall r\in(0,1)$，$T\overline{B}_1(0)\subset\dfrac{1}{r}T\overline{B}_r(0)\subset\overline{B}_{\beta/r}(0)\Rightarrow\|T\|\leqslant\beta/r$.

2. 令 $A = \mathrm{diag}(a_1, a_2, \cdots)$，由命题2.2.2有 $\|A\|_1 = \|A\|_\infty = \|a\|_\infty$.

3. 若 $T \in L(l^2)$，则 $|a_i| = \|Te_i\|_2 \leqslant \|T\|$，故 $a \in l^\infty$ 且 $\|a\|_\infty \leqslant \|T\|$. 若 $a \in l^\infty$，则 $\|Tx\|_2 \leqslant \|a\|_\infty \|x\|_2$，$\|T\| \leqslant \|a\|_\infty$.

4. $\|T_n\| = 1$.

5. 用 $\left|\sum_j a_{ij}x_j\right| \leqslant \left(\sum_j |a_{ij}|^q\right)^{1/q} \|x\|_p$ 推出 $\|Ax\|_p \leqslant \beta \|x\|_p$.

6. $\|T\| = \|a\|_\infty$. 在证 $\|T\| \geqslant \|a\|_\infty$ 时利用：$\forall \beta < \|a\|_\infty$，令 $A = \{x \in J : |a(x)| > \beta\}$，$u = \chi_A$，则 $\beta^2 mA \leqslant \|Tu\|_2^2 \leqslant \|T\|^2 mA$，$m$ 是 Lebesgue 测度.

7. 用命题2.2.3，$\|T\|_1 = 1/2$，$\|T\|_\infty = 1$.

8. 用命题2.2.3，$\|T\|_1 = \|T\|_\infty = 1/2$.

9. 用命题2.2.4，$\|T\| = 2/\pi$.

10. 用 Hölder 不等式，参考题5.

11. 用定理2.3.2，由 $\sum i^{-\alpha q} < \infty$ 得出 $1 \leqslant p < 1/(1-\alpha)$.

12. 用定理2.3.2，$\|u\| = \pi/\sqrt{6}$.

13. 用定理2.3.3，$\|f\| = \sqrt{\pi/2}$.

14. $\|f\| = 3/2$.

15. 用定理2.3.4，$\|f\| = 2$.

16. 用定理2.4.1.

17. 对 $T : A \times B \to X, (a,b) \mapsto a + b$ 用定理2.4.1.

18. 用定理2.4.2：若 $\{u_n\} \subset C(J)$，$u_n \rightrightarrows u$，$Tu_n \rightrightarrows v$，则必 $v = Tu$.

19. 用两次一致有界原理.

20. 只要证 $\beta \triangleq \sup_{\|x\|=1} \|Tx\| < \infty$；由一致有界原理，又只要对任给 $v \in Y^*$，指明 $\sup_{\|x\|=1} |v(Tx)| < \infty$.

21. 若 $0 \neq u \in A^\perp$，$a \in A$，则 $|u(x)|/\|u\| = |u(x-a)|/\|u\| \leqslant \|x - a\|$，这推出 $\beta \triangleq \sup\{|u(x)|/\|u\| : 0 \neq u \in A^\perp\} \leqslant d(x, A)$；用定理2.4.6证 $\beta \geqslant d(x, A)$.

22. 可设 A 是闭子空间，$u(x) = \langle x, a\rangle (x \in A)$，$v(x) = \langle x, b\rangle (x \in H)$，$a \in A$，$b \in H$，$\|a\| = \|b\|$，$u = v|A$，则必 $\|a - b\|^2 = 0$.

24. 证 $\overline{\mathrm{span}A} \subset {}^\perp(A^\perp)$ 是直接的. 若 $x \in (\overline{\mathrm{span}A})^c$，则由定理2.4.6有 $u \in A^\perp$，使 $u(x) \neq 0$.

25. 若 $\alpha(\cdot) : [a,b] \to \mathbf{K}$ 与 $x(\cdot) : [a,b] \to X$ 有连续导数，X 完备，则

$$\int_a^b \alpha'(t)x(t)\mathrm{d}t = \alpha(t)x(t)\Big|_a^b - \int_a^b \alpha(t)x'(t)\mathrm{d}t.$$

证法参照定理2.4.8.

26. 首先在 $\mathrm{span}\{x_i\}$ 上定义线性泛函 $u\left(\sum \beta_i x_i\right) = \sum \alpha_i \beta_i$，然后用定理2.4.5.

27. 若 $x_0 \in A^c$，则由定理2.4.14有 $u \in X^*$，$r \in \mathbf{R}$，使得 $u(x_0) < r < u(A)$，于是 $A \subset \{u \geqslant r\}$，$x_0 \notin \{u \geqslant r\}$.

28. 注意 $x_n \to x \Leftrightarrow \|x_n - x\|^2 \to 0$.

29. 用推论 2.4.7.

30. 注意 $|u_n(x_n) - u(x)| \leqslant |u_n(x_n) - u(x_n)| + |u(x_n) - u(x)|$.

31. 否则，有 $\{x_n\} \subset X$，使 $\|x_n\| = 1$，$\|Tx_n\| > n^2$. 令 $y_n = x_n/n$，则 $y_n \to 0$，而 $\{Ty_n\}$ 不可能弱收敛于零.

32. 参考定理 2.5.6 之证.

33. 参考例 1.3.9(i).

34. 首先对 $u = \chi_{[\alpha,\beta]}$ 证 $f_n(u) \to f(u)(n \to \infty)$，然后用定理 2.5.2.

35. 注意 $\|T_n\| \leqslant 1$，参考例 2.5.10.

第三章

1. 利用式(3.1.6).

2. 利用式(3.1.12).

3. 只要证 $r_\sigma(TS) \leqslant r_\sigma(ST)$，用谱半径公式，注意 $(TS)^n = T(ST)^{n-1}S$.

4. $\sigma(T) = \{0\} \Leftrightarrow r_\sigma(T) = 0 \Leftrightarrow \forall \lambda \in \mathbf{C}: \sum (\lambda T)^n$ 收敛(用定理 3.1.4).

5. 验证 $Tx = 0 \Rightarrow x = 0$；$\lambda \neq 0, Tx = \lambda x \Rightarrow x = 0$.

6. $Tx = \lambda x \Leftrightarrow x_i = \lambda^{i-1} x_1$；$N(T_\lambda) \neq \{0\} \Leftrightarrow |\lambda| < 1$，故 $\sigma_p(T) = D_1(0)$；$\|T\| \leqslant 1 \Rightarrow r_\sigma(T) \leqslant 1 \Rightarrow \sigma(T) = \overline{D_1(0)}$.

7. 取 σ 的可数稠子集 $\{a_i\}$，令 $Tx = (a_i x_i)$，则 $T \in L(l^1)$，$\sigma(T) = \sigma$.

8. $\lambda u(x) \equiv xu(x) \Rightarrow u(x) \equiv 0$，故 $\sigma_p(T) = \varnothing$；$\lambda \overline{\in} [0,1] \Rightarrow (\lambda - x)^{-1} u(x) \in C[0,1]$ $(u \in C[0,1])$，$\lambda \in [0,1] \Rightarrow (\lambda - x)^{-1} \overline{\in} C[0,1]$，故 $\sigma(T) = [0,1]$.

9. $\lambda \in \sigma(T) \Leftrightarrow \exists v \in C(J)$，使 $(\lambda - \varphi)^{-1} v \overline{\in} C(J) \Leftrightarrow \lambda \in \varphi(J)$.

10. $u = (I - \lambda T)^{-1} v = \sum_0^\infty \lambda^n T^n v$，$T$ 依例 3.1.5.

11. 验证 $f(T)(\lambda I - T) = (\lambda I - T)f(T) = I$.

12. 以 $T - S = R(\lambda, S)^{-1} - R(\lambda, T)^{-1}$ 代入.

13. 用谱映射定理与题 11.

14. 注意 $[\mu - (\lambda - \tau)^{-1}]^{-1} = \mu^{-1} + \mu^{-2}(\lambda - \mu^{-1} - \tau)^{-1}$，用定理 3.2.2 与题 11.

15. 注意 $f(\sigma(T)) = \sigma(f(T)) = \sigma(0) = \{0\}$.

16. $f(T)$ 可逆 $\Leftrightarrow 0 \overline{\in} \sigma(f(T)) = f(\sigma(T))$.

17. $f(\lambda) = \sqrt{\lambda}$ 在除去负半实轴的复平面上解析，令 $A = f(T)$，则 $A^2 = T$.

18. 令 $f(\lambda) = \ln\lambda$，类似于上题.

19. $Tx = \lambda x \Rightarrow x_{i+1} = i!\lambda^i x_1$，$\sigma(T) = \sigma_p(T) = \{0\}$.

20. $\sigma_p(T) = \{1/i : i \geqslant 1\}$，$\sigma(T) = \{0\} \cup \sigma_p(T)$.

21. 注意 $\sigma(T) = [0,1]$ 并用定理 3.3.6.

22. 令 $Tu(x) = \int_0^1 e^{x-y} u(y) dy$，则 $T \in CL(C[0,1])$，$\sigma(T) = \{0,1\}$，当 $\lambda \neq 1$ 时原方程有惟一解 $u \in C[0,1]$.

23. $f(\lambda) = \lambda g(\lambda), f(T) = Tg(T) \in \mathrm{CL}(X)$.

24. $x + T^*Tx = 0 \Rightarrow \|x\|^2 + \|Tx\|^2 = 0 \Rightarrow x = 0$

25. $x = Tx \Rightarrow \|x - T^*x\|^2 = 0$.

26. $y \in R(T)^{\perp} \Rightarrow m\|y\|^2 \leqslant \langle Ty, y\rangle = 0$, 故 $R(T) = H$.

27. 直接验证.

28. $\|T^*x\| = \|Tx\| \geqslant \beta\|x\| \Rightarrow T$ 与 T^* 均为单射且 $R(T)$ 为闭集，故 $R(T) = {}^{\perp}N(T^*) = H$, T 可逆.

29. 用 $N(T) = {}^{\perp}R(T^*)$.

30. 注意 $S^*T = I$.

31. 令 $\lambda = \alpha + \mathrm{i}\beta$, 则 $\|T_\lambda x\|^2 = \|T_\alpha x\|^2 + \beta^2\|x\|^2 \geqslant \beta^2\|x\|^2$.

32. 否则 $r_\sigma(T) = \|T\| = 0$.

33. $T^* = -T \Leftrightarrow (\mathrm{i}T)^* = \mathrm{i}T$.

34. $\sigma(T) \subset \mathrm{i}\mathbf{R}$.

35. $\langle Px, x\rangle = \|Px\|^2 = \langle P^*Px, x\rangle \Leftrightarrow P = P^*P \Leftrightarrow P = P^* = P^2$.

36. 注意 $\|x\|^2 = \|Px\|^2 + \|x - Px\|^2$.

37. T 必为闭算子(用定义 3.5.4).

38. 参照命题 3.5.9.

第四章

1. $Fx = \dfrac{\mathrm{d}}{\mathrm{d}t}F(tx)\Big|_{t=0} = F'(0)x$.

2. $F'(x) = \displaystyle\int_a^b f_x(t, x)\mathrm{d}x$.

3. $F'(u)h = \displaystyle\int_a^b f_x(t, u(t))h(t)\mathrm{d}t, u, h \in C(J)$.

4. 若 $f'(0) = \varphi, u_t(x) = tx$, 则应 $\dfrac{|f(u_t) - \varphi(u_t)|}{\|u_t\|_0} = \left|\dfrac{1}{2} - \dfrac{t}{|t|}\varphi(u_1)\right| \to 0(t \to 0)$, 此属不可能.

5. $T'(x,y)(h,k) = T(h,y) + T(x,k), T''(x,y)(h,k)(h_1,k_1) = T(h,k_1) + T(h_1,k), T'''(x,y) \equiv 0$.

6. $H''(x)hk = (G''(Fx)F'(x)h)F'(x)k + G'(Fx)(F''(x)hk)$.

7. $\|F(x,y) - F(\bar{x},\bar{y})\| \leqslant \mathrm{const}\|y - \bar{y}\| + \|F(x,y) - F(\bar{x},\bar{y})\|$ (用中值定理).

8. 由中值定理显然有 $\mathrm{Lip}F \leqslant \sup\limits_{x\in\Omega}\|F'(x)\|$. $\forall x \in \Omega, \varepsilon > 0$, 取 $\delta > 0$, 使当 $\|h\| < \delta$ 时 $\|F(x+h) - Fx - F'(x)h\| \leqslant \varepsilon\|h\|$, 因而 $\|F'(x)h\| \leqslant (\mathrm{Lip}F + \varepsilon)\|h\|$, 这推出 $\|F'(x)\| \leqslant \mathrm{Lip}F + \varepsilon$.

9. 用推论 2.4.7.

10. 对 $\varphi(t) = f(x + th)(0 \leqslant t \leqslant 1)$ 用 Taylor 公式.

11. 用 Taylor 公式: $\|F'(x)h\| = \|F(x+h) - Fx - \displaystyle\int_0^1(1-t)F''(x+th)h^2\mathrm{d}t\|$.

12. $\| Fx - Fy \| = \lim_n \| F_n x - F_n y \| \leqslant \overline{\lim_n}(\mathrm{Lip}F_n) \| x - y \|$.

13. 验证 $F\overline{B}_\rho(x_0) \subset \overline{B}_\rho(x_0)$.

14. $\| F^n x_0 - x^* \| \leqslant \sum_n^\infty (\mathrm{Lip}F^k) \| Fx_0 - x_0 \|$.

15. 若 $f(x) = \| x - Fx \|$ 在 x^* 取最小值,则必 $f(x^*) = 0$, 否则将有 $f(Fx^*) < f(x^*)$.

17. 令 $Fu = TGu + v$, T 是以 K 为核的积分算子, $Gu(x) = g(u(x))$, 则 $\mathrm{Lip}F \leqslant \| T \| \mathrm{Lip}G \leqslant \| T \| \mathrm{Lip}g < 1$.

18. 用归纳法.

19. 用链规则得 $(F^{-1})'(y)F'(x) = I, F'(x)(F^{-1})'(y) = I, y = Fx$.

20. 对 $G(x,y,z) \triangleq F(x,y) - z = 0$ 用隐函数定理.

21. 用命题 4.4.3 之证法.

22. 设 $\{x_n\} \subset D, (I - F)x_n \to x$, 不妨设 $Fx_n \to y$, 则 $x = (I - F)(x + y)$.

23. F 必为紧算子,因此可用 Schauder 不动点定理.

24. 不妨设 $Fx_n \to y, \lambda_n \to \lambda$, 必定 $\lambda > 0$, 因此 $x_n \to x = y/\lambda, Fx = \lambda x$.

25. 直接依不等式式(4.5.1)验证.

26. 利用 25 题.

27. 取定 $x, y \in X$, 令 $T = \{t \in [0,1] : f(t'x + ty) \leqslant t'f(x) + tf(y)\}$, 则 $0, 1/2, 1 \in T$; $\alpha, \beta \in T \Rightarrow (\alpha + \beta)/2 \in T$; T 是闭集,故必 $T = [0,1]$.

28. 若 $x \neq 0$, 则 $\partial \| x \|^2 = \{2u : u \in X^*, u(x) = \| x \|^2 = \| u \|^2\}$; $\partial \| x \|_{x=0} = \{0\}$.

29. 用 $u(x - y) \leqslant f(x) - f(y)$.

30. 验证 $\mathrm{D}f(x, h)$ 对 h 是次线性函数.

参考文献

1 Kreyszig E. 泛函分析引论及应用. 重庆:重庆出版社,1986

2 Schechter M. 泛函分析原理. 沈阳:辽宁科学技术出版社,1986

3 张恭庆,林源渠. 泛函分析讲义(上册). 北京:北京大学出版社,1987

4 夏道行,严绍宗. 实变函数与应用泛函分析基础. 上海:上海科学技术出版社,1987

5 陈景良. 近代分析数学概要. 北京:清华大学出版社,1987

6 张鸣歧. 应用泛函分析引论. 北京:北京理工大学出版社,1989

7 吴绍平. 泛函分析及其应用. 杭州:浙江大学出版社,1990

8 胡适耕. 实变函数. 北京:高等教育出版社,1999

9 王长清. 近代解析应用数学基础. 西安:西安电子科技大学出版社,2001

10 胡适耕. 泛函分析. 北京:高等教育出版社,2001

名词索引

名词按汉语拼音字母顺序排列. 名词后的数字指出该名词首次出现的页数; 当标出几个页数时, 表明该名词在几个不同的意义上使用.